

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 10977.4 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-2  
Perfect score: 4079  
Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09873409/runat\_27032003.115418.19228/app query fasta\_1.7544  
-DB=GenEmbl -QWTF=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=500 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 54887 @runat\_27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.hcg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4079	100.0	2856	6	AX339028 Sequence
2	4079	100.0	3177	6	AX339030 Sequence
3	4079	100.0	3621	6	AX339032 Sequence
4	4079	100.0	3702	6	AX339031 Sequence
5	3919	96.1	3699	6	AX478104 Sequence
6	3456	84.7	2066	6	AX339027 Sequence
7	2480	60.8	4018	5	GA9799
8	2476	60.7	3987	10	CRUGP1185
9	2476	60.7	4296	10	CRUGP1
10	2476	60.7	4304	10	CRUGP1165
11	2468	60.5	3858	4	OAU78609
12	2467.5	60.5	4084	10	MUSMDRA
13	2465.5	60.4	3924	6	AX024455
14	2465.5	60.4	3924	6	AX409652
15	2465.5	60.4	3924	9	HUMMDR3
16	2451	60.1	4927	10	AF257746
17	2442	59.9	4279	6	AX105057
18	2442	59.9	4279	6	AX105078
19	2442	59.9	4279	6	AX105080
20	2441	59.8	3934	4	CFA419568
21	2441	59.8	4045	12	AF269224
22	2441	59.8	4186	6	AX108654
23	2441	59.8	4195	6	AX108656
24	2441	59.8	4279	6	AX105082
25	2438	59.8	4317	4	AF045016
26	2438	59.8	4317	6	AX105059
27	2438	59.8	4323	10	AF286167
28	2437	59.7	2752	5	AF099732
29	2436	59.7	3840	6	AX481416
30	2436	59.7	3860	6	AX322787
31	2436	59.7	3860	6	AX322789
32	2436	59.7	3988	6	AX024454
33	2436	59.7	4192	9	AF016535
34	2436	59.7	4264	6	AR051647
35	2436	59.7	4264	6	AR051650
36	2436	59.7	4646	6	AX336420
37	2436	59.7	4646	6	AX336708
38	2436	59.7	4646	6	AX3391099
39	2436	59.7	4646	6	AX391099
40	2436	59.7	4646	9	HUMMDR1
41	2436	59.7	4669	6	I08557
42	2436	59.7	6505	6	AR028671
43	2436	59.7	8630	6	AX012320
44	2436	59.7	8630	6	AX012321
45	2436	59.7	9318	6	AR028672

ALIGNMENTS

RESULT 1

AX339028 AX339028 2856 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 10 from Patent WO0194400.  
DEFINITION  
ACCESSION AX339028  
VERSION AX339028.1 GI:18129120  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Frank, M.H. and Sayegh, M.H.  
AUTHORS  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001.  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source  
1..2856  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 888 a 540 c 652 g 776 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.87e-296 Length: 2856  
Score: 4079.00 Matches: 812  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-2 (1-812) x AX339028 (1-2856)  
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
Db 418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATATTCGAGACCATATTGGA 477  
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
Db 478 GTGGTGTAGTCAGAGAGCTGTTTGTTCGGGACCAACATCAGTAAACATATCAAGTATGGA 537  
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgAlaGluAlaAsnAlaTyrAsp 60  
Db 538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597  
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
Db 598 TTATATCATGGAGTTTCCTAATAAATTTAATACATTTGGTGGGGAAGAGAGCTCAAAATG 657  
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
Db 658 AGTGAGGCGCAGAACAGAGGATCGCAATTTGCTGCTGCTTGTAGTTGCAAAACCCAGATT 717  
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
Db 718 CTGATTTTAGATGAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777  
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140  
Db 778 GCACCTGGAGAGCGCAGCAAGGTCGACTACAACTCGTGGTGGTGGTGGTGGTGGTGGTGG 837  
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
Db 838 ATTCGAAGTCAGATTGATTTGTGACCTTAAGGATGGATGCTGCGGAGAGAGAGCA 897  
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
Db 898 CATGCTGAACCTAATGCCAAAACGAGGCTCTATATTACCTTGTGATGTCAAGGATATT 957  
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
Db 958 AAAAAAGCTGATGAACAGATGAGTCAATGACATATTTCTACTGAAAGAAAGACCAACTCA 1017

Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220  
Db 1018 CTTCCTCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1077  
Qy 221 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLeuLysIleIleLysLeuAsn 240  
Db 1078 ACCCAATCTAAGAGATAGTCTTCTCCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAC 1137  
Qy 241 LysProGluTTPProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260  
Db 1138 AAGCCTGAATGGCCTTTGTGGTCTCGGGACATTTGGCTTCTGTTCTAAATGGAACCTGTT 1197  
Qy 261 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnLys 280  
Db 1198 CATCCAGATTATTTCCATCATCTTTGCAAAAATTTATAACCATGTTTGGAAATAATGATAAA 1257  
Qy 281 ThrThrLeuLysHisAspAlaGluLysSerMetIlePheValIleIleGlyValIle 300  
Db 1258 ACCACATTTAAAGCATGATGCAGAAAATTTATTTCCATGATATTTCGTCTTAAATGGAACCTGTT 1317  
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320  
Db 1318 TGCITTTGTCAAGTTATTTTCAGGGGATTTTACGGCAGAGAGGGGAAATTTTAAAG 1377  
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340  
Db 1378 ATGAGATTAAGACACTTGGCCTTCAAGCCCATGTTATATCAGGATATTTCCTGGTTTGCAT 1437  
Qy 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360  
Db 1438 GAAAAGGAAAAACAGCACAGAGGCTTGACAACTAATATTAGCCATAGATATAGCACAAAT 1497  
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380  
Db 1498 CAAGAGACACAGGTTCCAGGATTCGCGTCTTAACACAAAATGCACATACATGAGACTT 1557  
Qy 381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400  
Db 1558 TCAGTTATCAITTCCTTTATATATGATGGATGGAGATGACATTCCTGATTCCTGATTTGCT 1617  
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420  
Db 1618 CCAGTACTTCCGCTGACAGGAATGATTTGAAACCCGACCAATGACTGGATTTGCCAACAAA 1677  
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440  
Db 1678 GATTAAGCAAGAACCTTAAGCATGCTGGAAGATAGCACTGAAGCTTTGGAGAAATATAGCT 1737  
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460  
Db 1738 ACTATAGTGTCAATTAACAAGGAAAAAGCCTTCGAGCAAAATGATGAAGAGATGCTTCAG 1797  
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480  
Db 1798 ACTCAACACAGAAAATACCTCGAAGAAAACACAGATTTATTGGAAGCTGTTATGCAATTCAGC 1857  
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValTyrLeuIleGln 500  
Db 1858 CATGCCITTTATATTTTGGCTTATGCATGACAGGCTTCGATTTGGAGCCTATTAAATTCAA 1917  
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520  
Db 1918 GCTGGAGCAATGACCCACAGAGGCGCATGTTCAATAGTTTTTACTGCAATTTGCATATGAGCT 1977  
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540  
Db 1978 ATGGCATCGGAAAAACGCTCGTTTGGCTCTCTGAATATTCAAAAGCCAAATTCGGGGCT 2037  
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560  
Db 2038 CGCATCTGTTTGGCTTGTGGAAAAAGAAACCAATATAGACAGCCGCGATCAAGAGGG 2097  
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580



Db 2098 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTCTATCCA 2157  
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600  
Db 2158 TGTGCGCCAGATGTTTTCATCTCGTGGCTATCCCTCAGTATTGAGCGAGGAAGACA 2217  
Qy 601 VallAlaPheValGlySerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArg 620  
Db 2218 GTAGCATTTGTGGGAGCAGCGCTGTGGGAAAGACACTTCTGTTCAACTTCGACAGA 2277  
Qy 621 LeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuLeu 640  
Db 2278 CTTTATGACCCCGTCAGAGCAAGAGTGCTGTTGTAGTGTGGATGCAAAAGAAATTGAAT 2337  
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660  
Db 2338 GTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCGTGTCTCTCAACTGC 2397  
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680  
Db 2398 AGCATTTGCTGAGAACATCGCTTATGTCACACAGCCGTGTGTCCTATAGATGAGATC 2457  
Qy 681 LysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700  
Db 2458 AAAGAAGCCGCAAAATGCAGCAAAATATCATCTCTTTATTGAAGGTCTCCCTGAGAAATAC 2517  
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720  
Db 2518 AACACACAAGTTGGACTGAAAGGACACAGCTTCTGCGGCGCAGAAACAAGACTAGCT 2577  
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740  
Db 2578 ATTGCAAGGGCTCTTCTCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2637  
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760  
Db 2638 CTCGATAATGACAGTGAAGGTGTTGTCAGCATGCCCTTGATTAAGCCAGGCGGGAAGG 2697  
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780  
Db 2698 ACATGCTAGTGGTCACTCACAGGCTCTCGCAATTCAGACGCGAGATTTGATGTTGTT 2757  
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800  
Db 2758 CTGCAATATGGAAGATAAAGAAACAAGAACTCATCAAGAGCTCTCGAGAAATCGAGAC 2817  
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812  
Db 2818 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 2  
AX339030 LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 12 from Patent WO0194400.  
ACCESSION AX339030  
VERSION AX339030.1 GI:18129122  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Frank, M.H. and Sayegh, M.H.

TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof

JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES Location/Qualifiers

source 1..3177

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 970 a 601 c 735 g 870 t 1 others

ORIGIN  
Alignment Scores:  
Pred. No.: 2,16e-296 Length: 3177  
Score: 4079.00 Matches: 812  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-2 (1-812) x AX339030 (1-3177)

Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
Db 739 ATGTGGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATTCGAGACCATATGGA 798  
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
Db 799 GTGGTTAGTCAAGAGCCCTGTTTTCGGGACCACCATCAGTAAACATATCAAGTATGGA 858  
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
Db 859 CGAGATGATGTGACTGATGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 918  
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
Db 919 TTTATCATGAGAGTTTCTTAATAAATTTAATACATTTGGTAGGGGAAAAAGGAGCTCAAAATG 978  
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
Db 979 AGTGGAGGCGAAGAACAGAGGATCGCAATTCCTGTCGCTTAGTTCCGAACCCCAAGATT 1038  
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
Db 1039 CTGATTTTAGATGAGGTACGTCCTGCCCTGGATTCAGAAAGCAAGTACGCTGTTCAGCT 1098  
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140  
Db 1099 GCACCTGGAAGAGGCGAGCAAAAGGTCCGACTCAATCTCGTGGTAGCAGCACCCGACTTTCTACT 1158  
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
Db 1159 ATTCCAAGTCAGATTTGATTTGTGACCTTAAGGATGAAGTCTGCGGAGAAAGAGCA 1218  
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
Db 1219 CATGCTGAATACTGGCAAAACGAGGCTATATATTACATTGTGATGTCACAGGATATT 1278  
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
Db 1279 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGAACCACTCA 1338  
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSer 220  
Db 1339 CTTCTCTCGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTTCAGCAAGGCTGAGGAATCC 1398  
Qy 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240  
Db 1399 ACCCAATCTAAGAGATGAAGTCTTCCTGGAAGTCTCTATTAAAAAATTTTAAAGTTAAAC 1458  
Qy 241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260  
Db 1459 AAGCCTGAATGGCCTTTTGTGTTCTGGGACATTTGGCTTCTGTTCTAAATGGAACCTGTT 1518  
Qy 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280  
Db 1519 CATCCAGTATTTTCCATCATCTTTTGCAAAAAATTTAACAACATGTTTGGAAATAATGATATA 1578  
Qy 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300  
Db 1579 ACCCATTTAAGCATGATGCAGAAATTTATTTCATGATATTTCGTCATTTTGGGTGTTATT 1638  
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLysLeuThr 320

```
Db 1639 TGGCTTTGTCAGTTATTTTCATGCGAGGATTTATTTACGAGACGAGGGAAATTTTAAACG 1698
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
Db 1699 ATGAGATTAAAGACACTTGGCTTCAAGCCATGTTATATCAGGATATTGCTGGTTTGTAT 1758
Qy 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 1759 GAAAGGAAACACGACAGGAGGCTTGACAAACAATATTAGCCATAGATATAGCACAAATT 1818
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 1819 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACAACAAATGCAACTAACATGGGACTT 1878
Qy 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
Db 1879 TCAGTTATCATTTTCCCTTTATATGATGGAGATGACATTTCTGATTTGATATTGCT 1938
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
Db 1939 CCAGTACTTGGCTGACAGGAATGATTGAACCGCAGCAATGACTGGATTTGCCAACAA 1998
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 1999 GATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2058
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGln 460
Db 2059 ACTATAGTGTCAATTAAACAGGGGAAAGGCTTCGAGCAAAATGTATGAAGAGATGCTTCAG 2118
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 2119 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGAGCTGTATGCATTCAGC 2178
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 2179 CATGCCCTTTATATTTTGGCTATGCGAGGAGGCTTCGATTTGGAGCTTATTTAAATCAA 2238
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 2239 GCTGGACGAATGACCCAGAGGCGATGTCATAGTTTTTACTGCAATTCGATGAGCT 2298
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2299 ATGGCCATCGGAAACGCTCGTTTGGCTCTCGAATATCCAAAGCCAAATCGGGGCT 2358
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2359 GGCATCTCTTTTGGCTTGTGGAAAGAAACCAAAATATAGACAGCGCAGTCAAGAAGGG 2418
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2419 AAAAGCCACAGACACATGTGAAGGAAATTTAGAGTTTCGAGAAGTCTCTTTCTCTATCCA 2478
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 2479 TCTCGCCACAGATTTTTCATCTCGTGGCTTATCCCTCAGTATTGAGCGAGAAAGACA 2538
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 2539 GTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACTTCTGTTCAACTTCGACAGA 2598
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 2599 CTTTATGACCCGTCAGAGCAAGTGTGTTGATGGTGTGGATGCAAAAGAAATTTGAAT 2658
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 2659 GTACAGTGGCTCGGTCCCAATAGCAATCGTTCCTCAAGACCTGTGCTCTTCAACTGC 2718
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
```

```
Db 2719 AGCATTGCTGAGAACATCGCCTATGTTGACAAACAGCCGCTGTTGCCATTAGATGAGATC 2778
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 2779 AAAGAAGCGGCAAAATGACAGCAAAATATCCATTTCTTTATTGAAGGTCTCCCTGAGAAATAC 2838
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAla 720
Db 2839 AACACACAAGTTGGACTGAAAGGAGCAGAGCTTTCTGCGCGCCAGAAACAAGACTAGCT 2898
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 2899 ATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTTATTGTTGGATGAGCCACTTCAGCC 2958
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 2959 CTCGATAATGACAGTGAAGGTGTTGAGATGCCCTTGATAAGCCAGGACGGGAAGG 3018
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3019 ACATGCTTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGTATGTTT 3078
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
Db 3079 CTGCACAATGGAAGATAAAGGAACAGGAACCTCATCAAGAGCTCTCGAGAAATCGAGAC 3138
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3139 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 3
LOCUS AX339032 3621 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 14 from Patent WO0194400.
ACCESSION AX339032
VERSION AX339032.1 GI:18129124
KEYWORDS human.
SOURCE ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .3621
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN

Alignment Scores:
Pred. No.: 2,57e-296 Length: 3621
Score: 4079.00 Matches: 812
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-873-409-2 (1-812) x AX339032 (1-3621)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1183 ATGTGGATGAGATGACATCAGAGCTTTAAATGTGGGCAATTATCGAGACCATATTGGA 1242
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1243 GTGGTTAGTCAAGAGCCCTTTTGTTCGGGACCACTCAGTAACAATATCAAGTATGGA 1302
Qy 41 ArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
```

Db 1303 CGAGATGATGCTGCTGATGAAAGAGATGAGAGAGCAGCAGCAAGGAGCAAAATGCGGTATGAT 1362  
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
Db 1363 TTTATCATGGAGTTTCCCTAAATAATTTAATACATTTGGTAGGGGAAAAAGGAGCTCAATG 1422  
Qy 81 SerGlyGlyGlnLysGlnArgileAlaileAlaAraGAlaLeuValAraAsnProLysile 100  
Db 1423 AGTGAGGAGCAGAAACAGAGGATCGCAATTTGCTGCTGCTTAGTTTCGAAACCCCAAGATT 1482  
Qy 101 LeuileLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValaGlnAla 120  
Db 1483 CTGATTTTAGATGAGCTACGCTCGCTCGCTGAGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542  
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrileValAlaAlaHieArgLeuSerThr 140  
Db 1543 GCATCGGAGAGGCGCAGCAAGGTCCGACTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1602  
Qy 141 IleArgSerAlaAspLeuileValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
Db 1603 ATTCGAAGTCAGATTGCTGATGACCTTAAGGATGGAATGCTGCGGAGAAAGAGCA 1662  
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
Db 1663 CATGCTGAACATAATGGCAAAACGAGGCTATATATTACATTTGCTGATGTCACAGGATATT 1722  
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
Db 1723 AAAAAGCTGATGAACAGATGAGTCAATGACATATTTCTCTGAAAGAAAGAACCAACTCA 1782  
Qy 201 LeuProLeuHisSerValLysSerileLysSerAspPheileAspLysAlaGluSer 220  
Db 1793 CTTCTCTGCACTCTGTGAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1842  
Qy 221 ThrGlnSerLysGluileSerLeuProGluValSerLeuLysLysLysLysLysLysLys 240  
Db 1843 ACCCAATCTAAGAGATAAGTCTTCTGAGGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902  
Qy 241 LysProGluTyrProPheValLysLysLysLysLysLysLysLysLysLysLysLysLys 260  
Db 1903 AAGCCTGAATGGCCTTTTGTGGTCTGAGGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1962  
Qy 261 HisProValPheSerilePheAlaLysLysLysLysLysLysLysLysLysLysLysLys 280  
Db 1963 CATCCAGTATTTTCCATCATCTTTGCAAAAATTAACCATGTTTGGAAATATGATATAA 2022  
Qy 281 ThrThrLeuLysHisAspAlaGluileTyrSerMetIlePheValileLeuGlyValile 300  
Db 2023 ACCACATTAAGCATGATGCAAAATTTATTCATGATATTCATGATATTCATGATATTCATGAT 2082  
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLysLysLysLys 320  
Db 2083 TGCTTTGTGCTGATTTTTCATGAGGATTTATTTATGAGGAGATTTATTTATGAGGAGATTT 2142  
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAsp 340  
Db 2143 ATGAGATTAAAGACATTTGGCCTTCAAGCCATGTTATATACAGATATTCGCTGGTTGAT 2202  
Qy 341 GluLysGluAsnSerThrGlyLeuThrThrileLeuAlaileAspIleAlaGlnile 360  
Db 2203 GAAAAGGAAAACAGCAGAGGCTTGACAACAATATTAGCCATAGATATAGCACAATTT 2262  
Qy 361 GlnGlyAlaThrGlySerArgileGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380  
Db 2263 CAAGGAGCAACAGGTTCCAGGATTTGCGCTTTAAACACAATAATGCAACTAACATGGGACTT 2322  
Qy 381 SerValileSerPheileTyrGlyTyrGluMetThrPheLeuileLeuSerileAla 400  
Db 2323 TCAGTTATCATTTCTTTATATGATGAGGAGATGACATTTCTGATTTCTGATTTGCT 2382  
Qy 401 ProValLeuAlaValThrGlyMetileGluThrAlaAlaMetThrGlyPheAlaAsnLys 420

Db 2383 CCAGTACTTCCGCTGACAGGANTGATTGAAACCCAGCAATGATGGATTGGCAACAAA 2442  
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysileAlaThrGluAlaLeuGluAsnileArg 440  
Db 2443 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAAATATACGT 2502  
Qy 441 ThrileValSerLeuThrArgGluLysAlaPheGluLysMetTyrGluGluMetLeuGln 460  
Db 2503 ACTATAGTGTCTATTAACAAGGAAAAAGCCCTTCGAGCAAAATGTATGAAGAGATGCTTCAG 2562  
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnileileGlySerCysTyrAlaPheSer 480  
Db 2563 ACTCAACACAGAAATACCTCGAAGAAAGCAGATATTGGAAGCTGTATGCATTCAGC 2622  
Qy 481 HisAlaPheileTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuileGln 500  
Db 2623 CATGCCITATATATTGCTCTATGCTATGACAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2682  
Qy 501 AlaGlyArgMetThrProGluGlyMetPheileValPheThrAlaileAlaTyrGlyAla 520  
Db 2683 GCTGGACGAATGACCCAGAGGCGATGTTTCATAGTTTTTACTGCAATTTGCATATGGAGCT 2742  
Qy 521 MetAlaileGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540  
Db 2743 ATGGCCATCGGAAAAAGCGCTGTTTGGCTCTCTGAATATTCGAAAGCCAAATCGGGGCT 2802  
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnileAspSerArgSerGlnGluGly 560  
Db 2803 CGCATCTGTTGCTTGTGTTGAAAGAAACCAATATAGACAGCCGAGTCAGAAAGGG 2862  
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580  
Db 2863 AAAAAGCAGACACATGTGAAGGGAATTTAGAGATTTTCGAGAAGTCTCTTTCTTCTATCCA 2922  
Qy 581 CysArgProAspValPheileLeuArgGlyLeuSerLeuSerileGluArgGlyLysThr 600  
Db 2923 TGTGCCCCAGATGTTTTCATCTCCGCTGCTTATCCCTCAGTATTTGACGAGGAAAGACA 2982  
Qy 601 ValAlaPheValGlySerSerGlyLysSerThrSerValGlnLeuLeuGlnArg 620  
Db 2983 GTAGCATTTTGGGGAGCAGCGCTGTGGGAAAGCACTTCTGTTCACTTCTGCAGAGA 3042  
Qy 621 LeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeuAsn 640  
Db 3043 CTTTATGACCCGCTGCAAGGACAAAGTGTGTTGATGCTGTGATGCGTGAAGAAATTTGAAT 3102  
Qy 641 ValGlnTyrLeuArgSerGlnileAlaileValProGlnGluProValLeuPheAsnCys 660  
Db 3103 GTACAGTGGCTCCGCTTCCCAATAGCAATCGTTCCTCAAGAGCCTGCTCTTCAACTGC 3162  
Qy 661 SerileAlaGluAsnileAlaTyrGlyAspAsnSerArgValProLeuAspGluile 680  
Db 3163 AGCATTTGCTGAGAACATCGCTATGGTGACAAACAGCCGTGTGGTCCCATTAGATGAGATC 3222  
Qy 681 LysGluAlaAlaAsnAlaAlaAsnileHisSerPheileGlyLeuProGluLysTyr 700  
Db 3223 AAAGAAGCGCAAAATGACGAAATATCCATTTCTTTTATTAAGAGTCTCCCTGAGAAATAC 3282  
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720  
Db 3283 AACACAAAGTTGGACTGAAAGGAGCAGCTTCTGCGGCGCAAGAAAGACTAGCT 3342  
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysLysLeuLeuLeuAspGluAlaThrSerAla 740  
Db 3343 ATTGCAGGGCTCTTCTCCAAAAACCCCAAAATTTATTTGTTGGATGAGGCCACTTCAGCC 3402  
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760  
Db 3403 CTCGATATGACAGTGAGAGGTGTTTCAGCATGCCCTTTGATAAGCCAGGACGGAGG 3462  
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaileGlnAsnAlaAspLeuileValVal 780  
Db 3463 ACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAGCCAGATTTGTATGTTGTT 3522

QY 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800  
 |||||  
 Db 3523 CTGCAAAATGGAAGATAAAGGAACCAAGGAACATCATCAAGAGCTCTGTGAAATCGAGAC 3582  
 |||||  
 QY 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812  
 |||||  
 Db 3583 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618  
 |||||

RESULT 4  
 LOCUS AX339031 3702 bp DNA linear PAT 09-JAN-2002  
 DEFINITION Sequence 13 from Patent WO0194400.  
 ACCESSION AX339031  
 VERSION AX339031.1 GI:18129123  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1  
 Frank, M.H. and Sayegh, M.H.  
 TITLE A gene encoding a multidrug resistance human p-glycoprotein  
 homologue on chromosome 7p15-21 and uses thereof  
 JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;  
 THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..3702  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 1098 a 708 c 856 g 1039 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,65e-296 Length: 3702  
 Score: 4079.00 Matches: 812  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-873-409-2 (1-812) x AX339031 (1-3702)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
 |||||  
 Db 1264 ATGGTGGATGGAATGACATCAGAGCTTTAAATGTGCGGCATATTCGAGACCATATTGGA 1323  
 |||||  
 QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
 |||||  
 Db 1324 GTGGTTAGTCAGAGCCTGTTTGTTCGGACCAACCATCAGTAACATATCAAGTATGGA 1383  
 |||||  
 QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
 |||||  
 Db 1384 CCAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGGGAAGCAATCGGTATGAT 1443  
 |||||  
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
 |||||  
 Db 1444 TTATCATGGAGTTTCTTAATAAATTTAATACATTTGGTGGGAAAAAGAGAGCTCAAAATG 1503  
 |||||  
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
 |||||  
 Db 1504 AGTGAGGCGAGAAACAGAGGATCGCAATTCCTGTCGCTTAGTTGGAACCCCAAGATT 1563  
 |||||  
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
 |||||  
 Db 1564 CTGATTTTAGTAGAGCTAGCTGCTGCTGCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623  
 |||||  
 QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140  
 |||||  
 Db 1624 GCACCTGGGAAGCGCAGCAAGCTCGGACTACAAATCGTGGTAGCACCGGACTTTCTACT 1683  
 |||||  
 QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
 |||||

Db 1684 ATTCGAAGTCAGATTTTGATTGTGACCCTAAAGGATGGAATCTCGCGGAGAAAGGACGA 1743  
 |||||  
 QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
 |||||  
 Db 1744 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTACATTGTGTGATGTACAGGATATT 1803  
 |||||  
 QY 181 LysLysAlaAspGluGlnMetClnSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
 |||||  
 Db 1804 AAAAAAGCTGATGAACAGATGAGTCAATGACATATTCTACTGAAGAAAGAACCAACTCA 1863  
 |||||  
 QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220  
 |||||  
 Db 1864 CTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAGGCTGAGGAATCC 1923  
 |||||  
 QY 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240  
 |||||  
 Db 1924 ACCCAATCTAAAGAGATAAGTCTTCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAAC 1983  
 |||||  
 QY 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260  
 |||||  
 Db 1984 AAGCCTGAATGGCCTTTTGTGTTCTGGGACATTTGGCTTCTGTTCTAAATGGAACCTGTT 2043  
 |||||  
 QY 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280  
 |||||  
 Db 2044 CATCCAGTATTTTCCATCATCTTTGCCAAAATTTATACCATGTTTGGAAATTAATGATAAA 2103  
 |||||  
 QY 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300  
 |||||  
 Db 2104 ACCCATTTAAAGCATGATGCAGAAATTTATTCCATGATATTCTGTTGGGTGTTATT 2163  
 |||||  
 QY 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320  
 |||||  
 Db 2164 TGCTTTGTGAGTTATTTCATGAGGGATTATTTTACGGCAGACGAGGGGAAATTTTAAAG 2223  
 |||||  
 QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAsp 340  
 |||||  
 Db 2224 ATGAGATTAAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCTGGTTGAT 2283  
 |||||  
 QY 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360  
 |||||  
 Db 2284 GAAAAGGAAAAACAGCACAGGAGGCTTGACAACTATATTAGCCATAGATATAGCACAAATT 2343  
 |||||  
 QY 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetClyLeu 380  
 |||||  
 Db 2344 CAAGAGCAACAGGTTCCAGGATTTGGCGCTTTAAACACAAAATGCAACTAACATGGGACTT 2403  
 |||||  
 QY 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400  
 |||||  
 Db 2404 TCAGTTATCATTTCTTTTATATGATGGATGGAGATGACATTCCTGATTTCTGAGTATTGCT 2463  
 |||||  
 QY 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420  
 |||||  
 Db 2464 CCAGTACTTGGCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTCGCCAACAA 2523  
 |||||  
 QY 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440  
 |||||  
 Db 2524 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTTGGAGATATACGT 2583  
 |||||  
 QY 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460  
 |||||  
 Db 2584 ACTATAGTGTCTTAACAAAGGGAAGAACCTTCGAGCAAAATGTATGAAGATGCTTCAG 2643  
 |||||  
 QY 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480  
 |||||  
 Db 2644 ACTCAACACAGAAATACCTCGAAGAAAGCAGACAGATTATTGGAAGCTGTTATGCAATTCAGC 2703  
 |||||  
 QY 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGln 500  
 |||||  
 Db 2704 CATGCTTTATATATTTTTCCTATGCGCAGAGGGTTTCGATTTGGAGCGCTATTATTAATCAA 2763  
 |||||  
 QY 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520  
 |||||  
 Db 2764 CTGGACGAAATGACCCCGAGGGCATGTTCTATAGTTTTTACTGCATTTGCAATGAGAGCT 2823  
 |||||

```
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2824 ATGGCCATCGGAAACACGCTGTTTGGCTCTCTGAATATTCACAAAGCCAAATCGGGGCT 2883
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnclucly 560
Db 2884 GCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGCCGAGTCAGAGGG 2943
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
Db 2944 AAAAGCCAGACACATGTGAAGGAATTTAGAGTTTCGAGAAGTCTCTTCTCTATCCA 3003
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 3004 TGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 3063
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 3064 GTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCTGTTCACACTTCTGCAGAGA 3123
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 3124 CTTATGACCCCGTCAGAGCAAGTGTCTTTGTATGTGTGGATGCAAGAAATTTGAAT 3183
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 3184 GTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCTCAAGAGCCGTGCTCTTCAACTGC 3243
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
Db 3244 AGCATGCTGAGAAACATGCCCTATGTGTGACACACGCGGTGTGTGCCATTAGATGAGATC 3303
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 3304 AAAGAGCCGCMAATGCAACAAATATCCATCTTTTATTGAAGGTCTCCTGAGAAATAC 3363
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3364 AACACACAAGTTGGACTGAAGGAGCAGACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3423
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3424 ATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTTATTGTTGGATGGAGCCACTTCAGCC 3483
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3484 CTCGATAATGACGTGAGAGGTGGTTGAGCATGCCCTTGATAAGCCAGGACGGGAAG 3543
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3544 ACATGCCCTAGTGTCTACACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGT 3603
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAsp 800
Db 3604 CTGCACAAATGGAAGATAAGGAACAAGGAATCATCAAGAGCTCTCGAGAAATCGAGAC 3663
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3664 ATATATTTAAGTTAGTGAATGACAGTCAGTGCAG 3699
RESULT 5
LOCUS AX478104 3699 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 34 from Patent WO0240541.
ACCESSION AX478104
VERSION AX478104.1 GI:22217064
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
```

```
AUTHORS Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
Walia,N.K., Yao,M.C., Baughn,M.R., Gandhi,A.R., Ding,L.,
Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
Asimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
Policky,J.L. and Kearney,L.
Transposers and ion channels.
Patent: WO 0240541-A 34 23-MAY-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1..3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="incyte ID No: 7472030CB1"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 2.62e-284 Length: 3699
Score: 3919.00 Matches: 785
Percent Similarity: 96.80% Conservative: 1
Best Local Similarity: 96.67% Mismatches: 0
Query Match: 96.08% Indels: 26
DB: 6 Gaps: 1
```

US-09-873-409-2 (1-812) x AX478104 (1-3699)

```
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1339 ATGTTGGATGAGATGATCATCGAGCTTTAAATGTGGCATTTATCGAGACCATATTGGA 1398
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1399 GTGGTTAGTCAAGAGCCTGTTTTCGGGACCACCATCAGTAAACAATATCAAGTATGGA 1458
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
Db 1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGCGACGCAAGGGAAGCAATGCGGTATGAT 1518
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMet 80
Db 1519 TTTATCATGGAGTTTCTTAATAATTTAATACATGTTGGTGGGGAAGAGAGCTCAATG 1578
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
Db 1579 AGTGAGGGCGAGAAACAGAGGATCGCAATTGCTGCTGCTTAGTTTCGAAACCCCAAGATT 1638
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 1639 CTGATTTTAGATGAGGCTACGCTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698
Qy 121 AlaLeuGluLysAlaSerLysGlyValArgThrThrIleValValAlaHisArgLeuSerThr 140
Db 1699 GCATCGGAAGAGCGAGCAAGGTCGGACTCAATCTGTTGGTAGCACCCGCTTTCTACT 1758
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
Db 1759 ATTGCAAGTCAGATTGATTGTGACCTTAAAGGATGGAATGCTGGCGGAGAAAGAGCG 1818
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
Db 1819 CATGCTGAACATAATGGCAAAACGAGGCTATATTATTACATTGTCATGTCAGGATATT 1878
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
Db 1879 AAAAAGCTGATGAACAGATCGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1938
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
Db 1939 CTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAGGCTGAGGAATCC 1998
Qy 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
```

```
Db 1999 ACCCAATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTAATAAATTTTAAAGTTAAAC 2058
Qy 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
Db 2059 AAGCCTGAATGACCTTTTGGTCTGGGGACATTTGGCTCTCTTAATGGAAGTGT 2118
Qy 261 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
Db 2119 CATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACCATGTTTGGAAATTAATGATAA 2178
Qy 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
Db 2179 ACCACATTAAGCATGATCAGAAATTTTATCCATGATATTCGTCTATTTGGGTGTTAT 2238
Qy 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
Db 2239 TGTCTTGTCTAGTTATTCATG----- 2259
Qy 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
Db 2260 -----CAGGATATTGCTTGGTTTGAT 2280
Qy 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
Db 2281 GAAAGGAAACAGACAGAGGCTTGACAACAATATTAGCCATAGATATAGCAAAAT 2340
Qy 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
Db 2341 CAAGGACCAACAGGTTCAGGATTCGGCTTTAAACAAATATGCAATCAATCGGACTT 2400
Qy 381 SerValIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400
Db 2401 TCAGTATCATCTTCTTTATATGATGGAGATGACATTCCTGATTCGAGTATTGCT 2460
Qy 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
Db 2461 CCAGTACTTGGCGTACAGGAATGATTTGAACCGCAGCAATGACTGGATTTGCCAACAAA 2520
Qy 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
Db 2521 GATAGCAGAACTTAAGCATGCTGGAAAGATAGCACTGAAGCTTTGGAGATATACGT 2580
Qy 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
Db 2581 ACTATAGTGTCTAATAACAGGGGAAAGGCTTCGAGCAAAATGTATGAGAGATGCTTCAG 2640
Qy 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
Db 2641 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATTTGGAAGCTGTATGCAATTCAGC 2700
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
Db 2701 CATGCTTTATATATTTCCTATGCGGAGGGTTTCGATTTGGAGCCTATTATTAATCCA 2760
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
Db 2761 GCTGACGAATACCCAGAGGCGCATGTTTCATAGTTTTACTGCAATTCATATGAGCT 2820
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
Db 2821 ATGGCCATCGGAACACGCTCGTTTGGCTCTCGAATATTTCAAAAGCCAAATTCGGGGCT 2880
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
Db 2881 GCGCATCTGTTTGGCTTTGTTGAAAGAAACCAATATATAGACAGCGCCAGTCAAGAGGG 2940
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrPro 580
Db 2941 AAAAGCCAGACACATGTCAAGGAAATTTAGAGTTTCGAGAGATCTCTTCTTCATCCA 3000
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
Db 3001 TGTGCGCCAGATGTTTTCATCTCTCGTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 3060
```

```
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 3061 GTAGCATTTTGGGGAGCAGCGGTGTGGAAAGACACTTCTGTTCAACTCTCTGCAGAGA 3120
Qy 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 3121 CTTTATGACCCCGTCAAGGACAAGTGTGTTGATGTGTGGATGCAAAAGAAATTGAAT 3180
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 3181 GTACAGTGGCTCCGTTCCAAATAGCAATCGTTCTCAAGAGCCTGTGCTCTTCAACTGC 3240
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
Db 3241 AGCATTGCTGAGAACATCGCTATGTCACAACAGCGGTGTGGTGCCATTAGATGAGATC 3300
Qy 681 LysGluAlaIleAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
Db 3301 AAAGAAGCGCAAAATGCAGCAATATCATCTCTTTTATTGAAGGTCTCTCCCTGAGAAATAC 3360
Qy 701 AsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3361 AACACACAAGTTGGACTGGAAGGACACAGCTTCTGCGCGCCAGAAACAAAGACTAGCT 3420
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3421 ATTTCAAGGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3480
Qy 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3481 CTCGATAATGACAGTGAGAAGGTGTTTCCAGCATGCGCTTGATAAAGCCAGGACGGGAAGG 3540
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3541 ACATCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCGAGATTTGATAGTGT 3600
Qy 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
Db 3601 CTGCACAATGAAAGATAAAGGAAACCAAGAACTCATCAAGAGCTCTCTGAGAAATCGAGAC 3660
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3661 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

```
RESULT 6
AX339027 LOCUS AX339027 2066 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 9 from Patent WO0194400.
ACCESSION AX339027
VERSION AX339027.1 GI:18129119
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
Patent: WO 0194400-A 9 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. 2066
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 647 a 403 c 470 g 546 t
ORIGIN
Alignment Scores: 6.27e-250 Length: 2066
Pred. No.: 3456.00 Matches: 687
Score: 100.00% Conservative: 0
```



Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 84.73% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-2 (1-812) x AX339027 (1-2066)

Qy 126 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 145  
Db 3 AGCAAGGTCGAGCTACATCGTGTGTAGCACACCGACTTCTTACTATTTCGAAGTCAGAT 62

Qy 146 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyValAlaHisAlaGluLeuMet 165  
Db 63 TTGATTGTGACCTTAAGAGATGGAATGCTGGCGGAGAAAGGAGACATGCTGCACTAATG 122

Qy 166 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 185  
Db 123 GCAAAACGAGGTCATATATTACTTGTGTGATGTACAGGATATTAAAAAGCTGATGAA 182

Qy 186 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 205  
Db 183 CAGATGGAGTCGAATGACATATTCTACTGAAGAGAAAGACCAACTCACTTCTCTGCACCT 242

Qy 206 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 225  
Db 243 GTGAAGAGCATCAAGTCAGACTTCATTGCAAGGCTGAGGAATCCACCCCAATCTAAAGAG 302

Qy 226 IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrPro 245  
Db 303 ATAAGTCTTCTGAGTCTCTCTATTAAAAATTTTAAAGTTAAACAGCTGAATGGCT 362

Qy 246 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 265  
Db 363 TTTGTGGTCTGGGACATGGCTTCTGTCTTAAATGGAATGTTTCACTCACTATTCTCC 422

Qy 266 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 285  
Db 423 ATCATCTTTGGCAAAATTAATACCATGTTTGGAAATAATGATAAAACCAATTAAGCAT 482

Qy 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305  
Db 483 GATGAGAAATTTATTCATGATATTCGTCAATTTGGTGTTATTGCTTGTCTGATAT 542

Qy 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325  
Db 543 TTTCAATGCGAGGATATTTTACGCGACAGCAGGGGAAATTTTAAACGATGAGATTAAGACAC 602

Qy 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer 345  
Db 603 TTTGGCTTCAAGCCATGTTATATCAGGATATTGCTGTTTGTGATAAAGGAAACAGC 662

Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365  
Db 663 ACAGAGGCTTGACAAATATATTAGCCATAGATATAGCAAAATTCAGGAGCAACAGGT 722

Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385  
Db 723 TCCAGATTGGCGTCTTAACACAAATGCACTTAACATGGGACTTTCAGTTATCATTTCC 782

Qy 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405  
Db 783 TTTATATATGATGGAGATGATCATTTCTGTGATTTGCTGATTTGCTCCAGTACTTCCCGTG 842

Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425  
Db 843 ACAGGAATGATTGAACCCGAGCAATGACTGGATTTTGCCCAAAAGATAAGCAAGACTT 902

Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445  
Db 903 AAGCATGCTGGAAGATACCACTGAAGCTTTGGAGATATACGACTATAGTGTCAITTA 962

Qy 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465  
Db 963 ACAAGGGAAAAAGCCTTCGAGCAAAATGTTAAGAGATGCTTTCAGACTCAACACAGAAAT 1022

Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485  
Db 1023 ACCTCGAAAGAACACACAGATTATTGGAAGCTGTTATGCAATTCAGCCATGCTTTATATAT 1082

Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505  
Db 1083 TTTGGCTATGAGCAGAGGTTTCGATTTTGGAGCCTATTAAATTCAGAGTGGAGCAATGACC 1142

Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525  
Db 1143 CCAGAGGCAATGTTTCATAGTTTTCATGCAATTCATATGAGCTATGCGCATCGGAAAA 1202

Qy 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545  
Db 1203 ACGTCTGTTTGGCTCTCTGAATATTCCAAAGCCAAATCCGGGGCTGCGCATCTGTTTGGC 1262

Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565  
Db 1263 TTTTGGAAAAAGAAACCAATATAGACAGCCGAGTCAAGAGGGGAAAGGCAACACACA 1322

Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585  
Db 1323 TGTGAAGGCAATTTAGAGTTTCGAGAGCTCTTCTTCTATCCATGTCGCCAGATGTT 1382

Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605  
Db 1383 TTTCACTCTCGTGGCTTATCCTCAGTATTGAGCGAGGAAAGACAGTAGTAGCATTTTGGGG 1442

Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625  
Db 1443 AGCAGCGCTGTGGAAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTG 1502

Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArg 645  
Db 1503 CAAGACAAAGTCTGTTGATGGTGGATGCAAAAGAAATTTGAATGTACAGTGGCTCCGT 1562

Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665  
Db 1563 TCCCAATAGCAATCGTTTCTCAAGAGCCTGTGCTCTTCACTGCAGCATTTGCTGAGAAC 1622

Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685  
Db 1623 ATCGCTATGTGTGACAAACAGCCGTGTGTGCTGATAGATGAGATCAAAAGAGCCCAAT 1682

Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705  
Db 1683 GCAGCAATATCCATCTTTTATTGAAGTCTTCCCTGAGAAATACACACACAGTTGA 1742

Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725  
Db 1743 CTGAAAGGAGCACACAGCTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGCTCTT 1802

Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745  
Db 1803 CTCCTAAAAACCCAAAAATTTTATTGTTGATGAGGCACTTCAGCCCTCGATAATACAGT 1862

Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765  
Db 1863 GAGAAGGTGGTTACAGATGCCCTTGATAAGCCAGGAGCGGAAGGACATGCCCTAGTGGTC 1922

Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785  
Db 1923 ACTCACAGCTCTCTGCAATTCAGAACCGCAGATTGTAGTGGTCTGCAACATCGAAAG 1982

Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeu 805  
Db 1983 ATAAAGGAACAAGGAACCTCATCAAGAGCTCTCGAGAAATCGAGACATATATTTAAGTTA 2042

Qy 806 ValAsnAlaGlnSerValGln 812  
Db 2043 GTGAATGCACAGTCAAGTCAG 2063

```
RESULT 7
GGA9799
LOCUS      Gallus gallus mRNA for ABC transporter protein. 4018 bp mRNA linear VRT 03-MAR-1999
DEFINITION
ACCESSION  AJ009799
VERSION     AJ009799.1 GI:3355756
KEYWORDS   ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE     Gallus gallus.
ORGANISM   Gallus gallus
REFERENCE  1 (bases 1 to 4018)
AUTHORS   Edelman, H.M., Duchek, P., Rosenthal, F.E., Foger, N., Glackin, C., Kane, S.E. and Kuchler, K.
TITLE     Cmr1, a chicken P-glycoprotein, confers multidrug resistance and interacts with estradiol
JOURNAL   Biol. Chem. 380 (2), 231-241 (1999)
MEDLINE   99209805
PUBMED    10195430
FEATURES   Location/Qualifiers
            1..4018
             /organism="Gallus gallus"
             /db_xref="taxon:9031"
             /tissue_type="intestine"
             /gene="cmdr1"
             /note="ABC transporter protein"
             /codon_start=1
             /product="P-glycoprotein"
             /protein_id="CAA08835.1"
             /db_xref="GI:3355757"
             /translation="MHSEDEKOKHTVDGNTVEIANSNODQDPDEDEKGGKKKKPKPM
             VSPLAFRYSCDTKLMLFGSLAIAHGTSPLIAMIIFGMDTDSFVTSGMTNIGNS
             SGNLSADYFNKLEEMTRYIYYSAIAAALVLAAYIQTISFWLTAAGQVKRIKFF
             HAIMEQETGFWNDAGELNTRLIDVSKINEIGDKIGLFIQSELTFLTGFIVGFIR
             GWKLTVLIVASPVGLSAALWAKILTFTDKEQAYAKAGAVEVLSAVRTVIAFG
             GQEKIKRYHKNLEDAKIRIGIKATISNIGMAFLIYASVALAFWYGTLLILANEY
             STGNLTVFFSLVIGAFSIGTAPSIKAFANARGAAYAIENIDNEPIDSVDAHKH
             PHIKGNLEFQNVFNPSRPDVEILKGLNLKYNCGQTVLVGSGCGSKSTVQLIQ
             FYDPKEGTTIDGDLKSLNVLYREIIGVNVQEPVLFAITTAENIRYRBDVTMEET
             ERATKEANAYDFIMKPKFTVVGEGQAQSGQKQRIATARALVHNPKILLDEAT
             SALDTSESVVQALDKAREGRTTVVVAHRLSTVRNADLIARVFSGVITEQGNHSLI
             EKKGIYKLVNMQTIETEDPSSEKSENAVSVKRSQSNLDESLLKELRRGSTRSMK
             KPEGNDEKSSPDEELPVPVSLKMLKNWEPYFVAGTFCAIUNGALQAPSVI
             FSEIIGTSETQKVRKSNLYSLFLALGLISFTFFVQSFAPFKAGEIUTMKLR
             MAFKAMLRQDMWAFDQKNSTLRLANDASQVKGATVGLALIQNLNLGII
             ISLVYQWLTLLVLAAPVPIIYVAGMIEMKMLAGHAKDKIELEAGKATAEIENRT
             VASLRKRYELMVEGHELLPVYRNVGVKXAHIPGFCFSLSQAMFFTYAGCFRFGAYLV
             VNGHTEYKTVLIVFSAVFGMALGQTSFPADYAKAKISAAHLFVLNRPVPPIDSVR
             EDGKPEFGGNTRIKDVKNPNRPEVKILQGLNLAVKEGTELALVSSGCGSKSTV
             QLLEFPYLPSEIVFDODIATKLNIQMLRSHIGVLSQRPILFDFTIAENIAYGNRSR
             EYVHEEIIISAAKAASIHGFIISLPEKYNTRVGDGTQISGGQKORIAIARALIRKPOI
             LLLDLETSALDTESEKIYQVEALDKAREGRTCTIVIAHRLSTIQNADKIAVIQNGKVIQ
             GTHQQLAEKFPYSLNVQSGCWN"
            112..3972
             /gene="cmdr1"
             /product="ABC transporter protein"
             /note="P-glycoprotein"
BASE COUNT  1288 a 758 c 911 g 1061 t
ORIGIN
Alignment Scores:
Pred. No.: 2,32e-176 Length: 4018
```

```
Score: 2480.00 Matches: 472
Percent Similarity: 75.70% Conservative: 151
Best Local Similarity: 57.35% Mismatches: 186
Query Match: 60.80% Indels: 14
DB: 5 Gaps: 1
US-09-873-409-2 (1-812) x GGA9799 (1-4018)
QY 2 ValAspGluAsnAspIleAArgAlaLeuAsnValAArgHisTyrAArgAspHisIleGlyVal 21
Db 1489 ATTGACGGCGAGGATCTGAAGAGCCTGAATGTAGATATCTCGAGAGATTATTGGTGTG 1548
QY 22 ValSerGlnGluProValIleuPheGlyThrThrIleSerAsnAsnIleLysTyrArg 41
Db 1549 GTGAACACGAGGAGCCGCTGCTTCTCAACTATTTGCAAAAAATATTGCTATGCGCCT 1608
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaAArgGluAlaAsnAlaTyrAspPhe 61
Db 1609 GAGGATGTCACCATCGAAGAAATTTGAAGAGCTACCAAGGAGCTAATGCTTATGATTTC 1668
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyValaGlnMetSer 81
Db 1669 ATCATGAAGCTCCCAAGAAGTTTGAAACTGTGTGTGAGAAAGAGGGCGCACAGATGAGT 1728
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1729 CGAGGGCGAAGCAGCAATAGCAATTCGCCGTGCTGCTGTTCACAACTCTAAATTTCTT 1788
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1789 CTGCTTGATGAGGCAACGTCAGCTTGGATCTAGAGATGGAATCCGTTGTGCGGCGAGCA 1848
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaAlaHisArgLeuSerThrIle 141
Db 1849 CTGGCAAGGCAAGAGAGGACGCGACCGTGTGTAGTAGCTCATCGTCTATCAACAGTC 1908
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1909 CCAATGTCAGATCTTATCGCTGTGTTGAAAGTGGAGTTATCACAGAAACAGAAATCAT 1968
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db 1969 TCTCACTGATTGAAAAAAGGAGATCTATATAACTTGTAACTACGACGACAAATAGAA 2028
QY 182 LysAlaAspGluGlnMetGlnSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db 2029 ACTGAAGATCCATCATCAGAAAAAAGTGAAATGCGAGTCAGTGTCAAAAAGAGTGATCT 2088
QY 202 ProLeuHisSerValLysSerIleLysSerAspPhe----- 213
Db 2089 CAATCAAAATCTGGATGAATCCCTAAAAAAGAAATTTGAGAAGAGGATCAACTAGGAGAAC 2148
QY 214 -----IleAspLysAlaGluGluSerThrGlnSerLysGluIleSer 227
Db 2149 ATGAAAAAGCCAGGAGAACCTTAATGATCTAGTGAAGAGGTAGTTTACCTGATGAGNA 2208
QY 228 LeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheVal 247
Db 2209 TTACCTCCAGTTTTCATTTCTGAAACTTATGAAGTTTGAACAAAAACGAATGGCCATCTTT 2268
QY 248 ValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIle 267
Db 2269 GTAGCTGGAACCTTCTGTGCAATTTGCAATGGAGCTTTACACACCTGCATTTTTCAGTCATA 2328
QY 268 PheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
Db 2329 TTCTCAGAAATTTATAGGATTTTTTCAGAAACTGATCAGAAAGTTTTCAGAGAAAAAGAGC 2388
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2389 AACTTATATTCACTGCTGTTTGTAGCACTGGGATCAATTCCTCTTTTACTTCTTCTTGT 2448
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
```



```

Db 2449 CAGGGTTTTCATTTGGCAAACTGGAGAAATCTTACAAATGAGCTTCGATTCAATGCA 2508
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTppPheAspGluLysGluAsnSerThrGly 347
Db 2509 TTTAAAGCAATCTGAGACAGACATGGCTGTTTGATGATCTCTAAATAATAGCACTGA 2568
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2569 GCATCTGACAAACAGACTTGTCTAATGATGCTCACAAGTAAAAAGGAGCAACTGGTCTCAGA 2628
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2629 TTGGCATTAAATGGCCAAACATAGCTAACTCTGGGAACCTGGAAATATTATATCTTAGTT 2688
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2689 TATGGCTGGCAGCTGACCTGTTACTCTTAGCTGTGTACCAATCATTCGACTGGCAGGA 2748
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2749 ATGATTGAAATGAAGATGTTGGCTGGCATGCAAAAAAGATAAAATAGAGCTGGAAGCT 2808
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAtqThrIleValSerLeuThrArg 447
Db 2809 CGAGGAAGATTTGCCACAGAGCCATAGAAAAATATTCGAACCTGTTGCTAGTTTGACTCGG 2868
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2869 GAAAAAGATTTGAGCTCATGTACGGAGAACATTTGCTGTGTACCATACAGAAATTCGTGTG 2928
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2929 AAGAAGGGCCACATATTTGGGTTTTGTTTTTCCCTTTCACAAGCAATGATGTTTTTACC 2988
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2989 TATGCTGGCTGTTCCGCTTTGGAGCCTATCTGTGGTTAAATGGACACATAGAGTACAAA 3048
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3049 ACTGTGTTTTAGTGTTCAGCTGTGTATTGTTGGTGTATGGCATTAGACAAACCAGC 3108
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3109 TCTTTTGGCCACAGACTATGCCAAAGCCAGATATCAGCAGCCCACTTGTGTGCTGTTT 3168
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3169 AATAGAGTACCTCCAATAGACAGTTACCCGCGAGGATGGAGAGAAACCTGAAAAATTTGGG 3228
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3229 GGAACACAGAGATCAAGATGTGAATTTAACTATCCAAACCGCAGAGGTCAAAATC 3288
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3289 CTCCAAGGCTTGAATCTAGCAGTAGAAAAAGGAGAAACGTTGGCCCTTGTGTAGCAGT 3348
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3349 GGATGTGGAAGAGCAGCTGTTGTCAGCTACTTGAGAGATTTTATGATCCACTCAGTGGA 3408
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3409 GAAATTTGTTTGATGACATGATGCAAGACACTAAATATATCAGTGGCTGAGATCTCAC 3468
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3469 ATCGGTATGCTCTCAAGAGCCAAATCCTGTTCCGACTTCACCATCGCTGAAAAACATCGG 3528
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAlaAla 687

```

```

Db 3529 TACGGGGACACAGCTCGGAGGTGTCCCATCAGGAAATCATTAGTCAGCGCAAGACGCC 3588
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrIleValGlyLeuLys 707
Db 3589 AGTATTTCATTCCTTCATTGACTCTCTGCCAGAAAAATACAATACTCGTGTAGGAGACAAG 3648
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaAtqAlaLeuLeuGln 727
Db 3649 GGAACACAGCTTTCTGGTGGTCAGAAAAACAGTATTGCTATTGCTCGAGCTCTTATACGT 3708
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3709 AAGCCCCAGATTCTGCTACTGATGAGCTACATCTGCTTGGACACAGAAAGCGAAAAAG 3768
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3769 ATTGTCCAGGAAGCACTGCGATAAAAGCCAGAGAAAGTGCACCTGCATCGTATAGCTCAC 3828
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3829 CGCCTGTCCACCATCCAGAAATGCTGACAAGATTGCTGTGATCCAGAACCGCAAGGTCTATA 3888
Qy 788 GluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3889 GAGCAGAGGACTCACCAGCAGCTCTGCTGCTGAAAAAGGCTTCTACTACTCGTGGTCAAT 3948
Qy 808 AlaGlnSer 810
Db 3949 GTTCAAGCT 3957

RESULT 8
CRUPGFI185 3987 bp mRNA linear ROD 27-APR-1993
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
DEFINITION
ACCESSION M59254
VERSION M59254.1 GI:191156
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone ADX185.
ORGANISM Cricetulus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE
1 (bases 1 to 3987)
AUTHORS Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
TITLE Full length and alternatively spliced pgpl transcripts in
MULTIDRUG-RESISTANT CHINESE HAMSTER LUNG CELLS
J. Biol. Chem. 266 (7), 4545-4555 (1991)
JOURNAL 91154265
MEDLINE 1671863
PUBMED
FEATURES
Location/Qualifiers
1..3987
/organism="Cricetulus sp."
/db_xref="taxon:10031"
/cell_type="fibroblast"
/tissue_type="lung"
1..3987
/gene="pgp-1"
115..3624
/gene="pgp-1"
/codon_start=1
/product="p-glycoprotein"
/protein_id="AAA37005.1"
/db_xref="GI:191157"
translation="MTTYAYVYTGAGVLIVAYIQVSWFCAAAGROIHKIRKQFFHA
IMNQEIGWFDVHDVGEINLTRLTDDVSKINSGIDKIGMFFQAMATFGGFTIGTRGW
KLTLVILAISFVLGSLAGIKWAKILSSFTDELQYAKAGAVAEVLAIRVIAFGGQ
KKELERNYNNLEEARGLIKKAITANISMGAAPFLIIYASALAFWYGLSVLSKEYSI
GOVLTVFAVLIAFPFGQASPNTEAFANAGAYEIFNIDNPKSIDSFSKNGKPD
NIGKLEFKNIHFSPSRKDVQILKGLNLVQSQGTVALVGNSCGKSTTVQLQRLY
DPTEGVVSDIQDITINVLREIIGVWSQEPVLFTATTIENIRYGRNVMTBEIEK
AVKENAYDTIMKLPKHFEDTLVGERGAQLSGGQKORIAAPALVRNPKILLDEATSA
LDTESEAVQALDKAREGRTTVIAHRLSTVRNADIAGDGGVIVQEGNHSELMRE
KGIYFKLVMTQTAGNEVGEKNEIDNLDMSKDSASSLIRRRSTRRSIRGRPH

```

DDRLKSTKEALDEVPPIISFWRIILKNSSEWPYFVVGIFCAIVNGLQAPFISFK  
VVGVRTRNDSTKRDHNSLFLSLILGLVIFSTFFLQGGTFRGQNGILKRLRYMV  
FKSMRLQVDFPNKNTGALTTLRLANDAGQVKGATGARLAVITRQNLIGTGIIS  
LYGQQLTULLLAIVIPITAGVBMKMLSGOALKKEKLEGGSKIAATEAENPRTVV  
SLTRQKPNMYAQSLQIPRYNALKKAHVFGITFSFTQAMYPFSAACFRGALVAR  
ELMTFENVLLVLSAIVGMAVQGVSSFAPDYAKAKVSGSHIIMIERKVPISDYSVG  
GLKPNTEGNKFNVEVNYPRPDI PVLQGLNLEKVGQTLALVSGGCGKSTVQL  
LERFYDPWAGTVFLDGKEVNLQVWLRHLGVSEPTLFDSCIAENIAYGDNRSVV  
SODELRADKEANIHQFTESLPDKYNTRVGDGKTLGSGQKORIAIALRALVROPHILL  
LDEANSALDTESEKVVQREALKARSGRTCIIVIAHRLSTIQNADLLIVVINGKVBHGT  
HQOLLAQKGIYFSMVSVQAKR"

BASE COUNT 1194 a 822 c 944 g 1027 t  
ORIGIN Map position 1q26.

## Alignment Scores:

Pred. No.: 4 58e-176 Length: 3987  
Score: 2476.00 Matches: 468  
Percent Similarity: 77.52% Conservative: 170  
Best Local Similarity: 56.87% Mismatches: 171  
Query Match: 60.70% Indels: 14  
DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x CRUPGP1185 (1-3987)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21  
Db 1141 ATCGACGGACAGACATCAGGACCAATGTGAGGTATCTGCGGGAATCATTTGGGTC 1200  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIysTyrGlyArg 41  
Db 1201 GTGAGTCAGGAACCTGTGTTTGGCACCATCTATCGCTGAAACATTCGCTATGGCCGA 1260  
Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe 61  
Db 1261 GAAATGTCTACTATGACGAGATCGGAAGCTGTCAAGAACCAATCGCTATGACTTC 1320  
Qy 62 IleMetGluPheProAsnIysPheAsnThrLeuValGlyGluIysGlyAlaGlnMetSer 81  
Db 1321 ATCATGAAACTGCCCATAAATTTGACACTCTTGTGTTGGTGGAGAGGAGCACAGCTGAGT 1380  
Qy 82 GlyGlyGlnIysGlnArgIleAlaAlaArgAlaLeuValArgAsnProIysIleLeu 101  
Db 1381 GGAGGACAGAAACAGAGAATCGCATTTGCTGTCGCCCTGCTCGCAACCCCAAGATCCTT 1440  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAlaA 121  
Db 1441 TTGTTGATGAGCGGACATCAGCCTTGACACAGAAGTGAAGCCGTGTTACAGGTGCT 1500  
Qy 122 LeuGluIysAlaSerIysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 1501 CTGGATAAGCTAGAGAGGCGGACTACCATTTGTATAGCCACCCGCTTGTCTACAGTT 1560  
Qy 142 ArgSerAlaAspLeuIleValThrLeuIysAspGlyMetLeuAlaGlnIysGlyAlaHis 161  
Db 1561 CGAAATGCTGACATCATAGCTGCGGTTTGATGTGTGTCATTTGTGGAGCAAGAAATCAT 1620  
Qy 162 AlaGluLeuMetAlaIysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
Db 1621 GAAGAGCTCATGAGAGAGAGGCGCATTTACTCAAACTTGTCTATGACACAGACAGAGA 1680  
Qy 179 -----AspIleLysLeuAlaAspGluGlnMetGluSerMetThr 191  
Db 1681 AATGAAATTTGAATAGGAATGAAGTTGGTGTGACTCTAAATAATGAAATTTGATATTAGAC 1740  
Qy 192 TyrSerThrGluArgIysThrAsnSerLeu-----ProLeuHisSerValIysSerIle 209  
Db 1741 ATGCTCTCAAAAGATTACGATCATCGAGTCTAATCAGAAGGAGATCAACTCGCAGAGATATC 1800  
Qy 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228  
Db 1801 CGTGGACCATGACCAAGACAGAGAGCTTAGTACCAGAGGCGCTTGGATGAAGATGTA 1860  
Qy 229 ProGluValSerLeuLeuIysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248

Db 1861 CCTCCAATTTCTCTTTGGCGTATCTCTAAAGTTGAATTCATCTGAATGGCCATATTTTGG 1920  
Qy 249 LeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIleIlePhe 268  
Db 1921 GTTGTATATCTGTGCCATGTAATAATGAGGCTTGCAACACGACATTTCTCAATATATTT 1980  
Qy 269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287  
Db 1981 TCTAAGTTGTAGGGGTTTTCACAGAAATATCTGATGATGTAAGCAACCAACGACATGATAC 2040  
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
Db 2041 AACTTGTCTTCTTATGTTTCTGATCTCTGGGTCATTTCTTTTATACATTTTCTT 2100  
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
Db 2101 CAGGGCTTCACATTTTGGCAAGCTGGAGATCTCTCAAGTTAAAGGGGCTACAGAGCCAGA 2280  
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
Db 2161 TTCAATCCATGCTGAGACAGGATGTGAGCTGTTGATTAACCTTAAACCAACCACTGGA 2220  
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
Db 2221 GCATTCAGCACCCAGGCTGCCCAACGATGCTGTCAAGTTAAAGGGGCTACAGAGCCAGA 2280  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 2281 CTGCTGTCTATACCAGAAATATAGCAATCTTGGACAGGAATCATATATCCCTAATC 2340  
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2341 TATGCTGGCAGCTTCACACTTCTACTCTTAGCAATGTGCCCATCATTCGCAATAGCAGA 2400  
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
Db 2401 GTGTTGATGATGAAAATGTTGTCTGCAAGCAAGCACTAAAGATAAGAGGAGCTAGAA 2460  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIleValSerLeuThrArg 447  
Db 2461 TCTGGAAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCTCTTTTGTGACTCG 2520  
Qy 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
Db 2521 GAGCAGACTTTGAAATATGTATGCCAGAGCTTGCAGATACCATACAGAAATGCTCTG 2580  
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 2581 AAGAAAGCACACGCTCTTTGGGATCATTCTCTTCCACCGAGCCATGATGATATTTTCG 2640  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 2641 TATGCTGTCTTGTTCGGTTTGGTCTTACTTGGTGGCAGCGGAATTAATGACATTTGAA 2700  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
Db 2701 AATGTTCTATTAGTATCTCAGCTATTGCTTGTGTGCTTGCAGTGGCAGGTCAGT 2760  
Qy 528 ValLeuAlaProGluTyrSerLysAlaIysSerGlyAlaAlaHisLeuPheAlaLeu 547  
Db 2761 TCATTTGCTCTGACTATGCCAAAGCAAAAGTGTCCGATCGCATCATCATGATCAT 2820  
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGlu 567  
Db 2821 GAAAGTCCCTTCCATTGACAGCTACAGACCGAGGCGCTGAAGCCCTAATACATTTGAA 2880  
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
Db 2881 GGAATGTGAATTTAATGAAGTCGTGTTCAACTACCCACCCGCGCTGACATCCAGTG 2940  
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607

```

Db 2941 CTTGAGGGCTGAACCTGAGAGTGAAGAGGGCCAGACGCTGGCCCTGGTGGCGCAGCAGT 3000
Qy GlyCysGlySerThrSerValGlnLeuLeuGlnArgLeuTyAspProValGlnGly 627
Db 3001 GCGTCGGGAGAGCAGCAGTGTCTGAGCTGTCTGAGCGGTCTTACGACCCCTGCGCGCA 3060
Qy GlnValLeuPheAspGlyValAspAlaLysGluLeuLeuValClnTrpLeuArgSerGln 647
Db 3061 ACAGTGTCTTCTAGATGGCAAGAAAGTAAACCAACTAAATGTCTCAGTGGCTCGAGCACAC 3120
Qy IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3121 CTAGGCAATGTCTCCAGAGCCCATCTGTTTGACTGAGCATGCTGAGAACATCGCT 3180
Qy TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAlaAla 687
Db 3181 TATGAGACACAGCGGGTCTGTGTCAGGATGAGATCGAGAGGCGCGGAAGAGCC 3240
Qy AsnIleHisSerPheIleGluGlyLeuProGluLeuTyAsnThrClnValGlyLeuLys 707
Db 3241 AACATCCACCACTTCACTTGAAGTCACTGCTGACCAATACCAACACAGAGTAGGAGACAA 3300
Qy GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3301 GGGACTCAGCTGCGGGTGGGAGAACACGATCGCCATAGCAGCCCTCTGTCAGA 3360
Qy LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3361 CAGCTCACATTTGCTCTGTGATGAGCAACATCAGCCCTGGATACAGAGATGAAGAG 3420
Qy ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3421 GTTGTCCAGGAAGCTCTGACAAAGCCAGAGAAGCGCCACCTGCAATTTGATCGCTCAC 3480
Qy ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisnGlyLysIleLys 787
Db 3481 CGCCTGTCCACCATCCAGAACCGAGACTTGATCGTGGTGATTCAGAAATGGCAAGGTCAAG 3540
Qy GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyPhyLysLeuValAsn 807
Db 3541 GAGCATGGCACCACGAGCTGTGGCACAGAAAGGCATCTATTCTCCATGCTCAGT 3600
Qy AlaGlnSer 810
Db 3601 GTGCGGCT 3609

RESULT 9
CRUPGPI 4296 bp mRNA linear ROD 22-JUN-1995
LOCUS C.griseus p-glycoprotein (isoform class I) mRNA, complete cds.
DEFINITION M60040
ACCESSION M60040
VERSION M60040.1 GI:191164
KEYWORDS multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE C.griseus adult liver and ovary, cDNA to mRNA.
ORGANISM Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 4296)
Endicott,J.A., Sarangi,F. and Ling,V.
Complete cDNA sequences encoding the Chinese hamster p-glycoprotein
gene family
JOURNAL DNA Seq 2 (2), 89-101 (1991)
MEDLINE 92135886
PUBMED 1685679
FEATURES
    Location/Qualifiers
        1..4296
            /organism="Cricetulus griseus"
            /db_xref="taxon:10029"
            /tissue_type="liver and ovary"
            /dev_stage="adult"
            103..3933
            /note="putative"

```

CDS

```

/codon_start=1
/product="p-glycoprotein isoform I"
/protein_id="AAA68883.1"
/db_xref="GI:191165"
/translation="MEPEEDFSGRKDNFLKMGKSKKEKKEKPPVSVFTMFYAGH
LDRLMLVGLTAAIIGHVALPVLMLVFGDMTDSFASVGNITPTNATNATONVNSDIFG
KLEEMTYAYVYTGICAGVLIVAMIOVSFWCLAAAGROIHKIROKHFHATIMNOBIFG
DVHDVGLNRLTDDVSKINEGDKLGMFQAMATPFGGFLIGTRGKMLTLVLA
NLEBKSAGIWKILSSFTDKELQYAKAGAAVEELAAIKTVIAPGQKQLBRYNN
NLEAKRLGKKAITANISMAAPLLIYASALAFWYGTSLVTSKEYSIGQVLTFFA
NLIAGFISGASPNIEAFANAGAAVEIFNIIDNKPSIDSKNGYKPDNIKLEPK
NIHFSYPSKDVQIILKGLNKVSGQTVLVGNSCGKSTTVQLQLLYDPTEGVVSI
DGDITINRYLREIIGVVSQEPVLPATTIAENIRYGRNVMTDETEKAVKEANAYD
FIMKLPKFTLGERCAQLSGQKQRIARALVRNPKILLLEDEATSDLTSEAVV
QAALDKAREGRTTIIVIAHRLSTVNADIAGFDGVIQEQNHELMERKGIYKPLVM
TQAGNEIELGNEVESKEINIDNMSKDSFASLIRRRSTRIRSGHDDNKLSTK
EALDEVPPIFSFWRILKLNSEWYFVYGFCAIVNGALQAPAFSIIFSKVGVFTNRT
DDETKRHSNLSLLFLILISPTIFLOQFTGKAGEILTKRLRYVFKSMRLQDV
SWFDPKNTTGALTTRLANDAGQVKGATGARLAVITONIANLGTGIIISLYGHQTL
LLLAIPILAIAGVVENKMLSGQALKDKKELEGSGKIATEAENFRVVSILTRQKPE
NMAQSLQIPIRNALKKAHVFGITFSFTQAMMYFYAACFRFGAYLVARELMTFENVL
LVFSALVFGMAVQGVSSFAPYAKAKVASHIIMIIEKVPISDSYTGGLKPLNTEG
NVKEVNVFNYPTRPDIPLVQLNLNLEVKGTGLALVGSNGCGKSTTVQLLFPDPA
GTVPLDGKENVOLNVMRLAHLGTVSOEPIILFDCSIAENIAGNSVNSWODETERAA
KEANIHQFISLDPKYNTRVGDGKTQLSGQKQRIARALVRNPKILLLEDEATSDLT
TESKVVQEAALDKAREGRTCIIVIAHRLSTQNADLIIVIQNGKVEHGHQQLLAQKG
IYFSMVSVOAGAKR"
BASE COUNT 1271 a 892 c 1029 g 1104 t
ORIGIN
Alignment Scores:
Prgm. No.: 5.06e-176 Length: 4296
Score: 2476.00 Matches: 468
Percent Similarity: 77.52% Conservative: 170
Best Local Similarity: 56.87% Mismatches: 171
Query Match: 60.70% Indels: 14
Db: 4
US-09-873-409-2 (1-812) x CRUPGPI (1-4296)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1450 ATCGAGCGTCAGACATCAGGACCATCATGTGAGTATCTCGGGAATCATTTGGGTC 1509
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1510 GTGAGTCAGAACCTGTGTTTTCGCCACCATCTATCGCTGAAAAACATTCGCTATGCGCGA 1569
Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1570 GAAATGTCTATGACGAGATCGAGAAAGCTGTCAAGGAAGCAATGCTTATGACTTC 1629
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluAlaGlnMetSer 81
Db 1630 ATCATGAACCTCCCAATAATTGACACTCTTGTGTGAGAGAGGACACAGCTGAGT 1689
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1690 GGAGCAGACAAACAGAAATCGCCATTGCTGCGCTGTCGCAACCCCAAGATCCTT 1749
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1750 TTGTTGGATGAGCGACATCAGCTTGGACACAGAAAGTGAAGCCGTGTTTCAGGCTGT 1809
Qy 122 LeuGluLysAlaSerLysGlyArgThrIleValAlaAlaHisArgLeuSerThrIle 141
Db 1810 CTGGATAGGCTAGAGAGCGGACCTACCATTTGTATAGACCCAGCTGTGTCTACAGTT 1869
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1870 CGAATGCTGCATCATCATGCTGGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1929
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyTyrSerLeuValMetSerGln----- 178

```

Db 1930 GAAGAGCTCATGAGAGAGGCGCATTTCAAACTTGTGCATGACACAGACAGCAGGA 1989  
Qy 179 -----AspIleLysAlaAspGluGlnMetGluSerMetThr 191  
Db 1990 AATGAATGAATTAGGAATGAAGTTGGTGTCTAAATAATGAATTTAGAC 2049  
Qy 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209  
Db 2050 ATGCTCTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACTCCGAGAGTATC 2109  
Qy 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228  
Db 2110 CGTGGACCATGACCAAGACAGAGAGCTTAGTACCAGAGAGGCGCTTGGATGAAGATGA 2169  
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluThrProPheValVal 248  
Db 2170 CCTCCAATTTCCTTTGGCGTATCCTAAGTTCATCTCAATGGCCATATTTTGTG 2229  
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
Db 2230 GTTGGTATATCTGTGCCATAGTAATGGAGCCTTGCAACCCAGCATTTCTCAATAATATTT 2289  
Qy 269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287  
Db 2290 TCTAAGGTTGTAGGGGTTTTCACAGAAATACTGATGATGAACCAACGACATGATAGC 2349  
Qy 288 GluIleThrSerMetIlePheValIleLeuGlyValIleCysPheValSerThrPheMet 307  
Db 2350 AACTTGTTCCTTATGTTCTGATCCTTGGGTCATTTCTTTATACATTTTCTT 2409  
Qy 308 GlnGlyLeuPheThrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
Db 2410 CAGGGCTTCACATTTGGCAAGCTGAGAGATCTCACAAGGCGACTCCGATACATGTT 2469  
Qy 328 PheLysAlaMetLeuThrGlnAspIleAlaThrPheAspGluLysGluAsnSerThrGly 347  
Db 2470 TTCATATCCATGTCAGACAGAGTGTGAGCTGTTGTATTAACCTTAAACACCACTGGA 2529  
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
Db 2530 GCATTGCCACACAGGCTCGCAACGATGCTGGTCAAGTTAAAGGGCTACAGGAGCCAGA 2589  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 2590 CTTGCTGTCTATACCCAGATATAGCAAAATCTTGGACAGGAATCATATCCCTAATC 2649  
Qy 388 TyrGlyThrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2650 TATGGCTGGCAGTTGACATTCCTACTCTTAGCAATTTGTCCTCCCATCATTTGCAATAGCAGGA 2709  
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
Db 2710 GTGGTTGAGATGAATGTTGCTCGCAACGACTAAAGATGAAGAGGCTAGAAGGT 2769  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
Db 2770 TCTGGAAGATTGCTACTGAAGCAATAGAGAACTTCGCACTGCTCTTTTGAATCTCG 2829  
Qy 448 GluLysAlaPheGluGlnMetThrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
Db 2830 GAGCAAGATTGCAAAATATGATGTCAGAGGCTTGACATACCATACAGAAATGCTCTG 2889  
Qy 468 LysLysAlaGlnIleIleGlySerCysThrAlaPheSerHisAlaPheIleThrPheAla 487  
Db 2890 AAGAAAGCACACGCTTTGGGATGCATCTCTCTTACCAGGCGCATGATGATTTTTCG 2949  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaThrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 2950 TATGCTGCTGTTTCCGTTTGGTGTCTTACTTGTGGCAGCGGAATTAATGACATTTGAA 3009  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527

Db 3010 AATGTTCTATTAGTATTCTCAGCTATTGCTTTTGGTGCATGGCAGTTGGCGAGGTCAGT 3069  
Qy 528 ValLeuAlaProGluLysSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
Db 3070 TCATTGCTCTCTGACTATGCCAAAGCCAAAGTGTGGCATCGCATATCATGATCATTT 3129  
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
Db 3130 GAAAGAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCGCTGAAGCCCTAATACATTTGAA 3189  
Qy 568 GlyAsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPheIle 587  
Db 3190 GGAATGTGAATTTAATGAAGTCTGTTCACATCCACCACCGGCTGACATCCAGTG 3249  
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607  
Db 3250 CTTAGGGGCTGAACCTGAGGTGAAGAGGCGCAGACGCTGGCCCTGTGTGGCAGCAGT 3309  
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuThrAspProValGlnGly 627  
Db 3310 GGCTCGGGAAGACACATGGTGTCCAGCTGTCTGAGCGGTTCTACGACCCCATGGCCGA 3369  
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGln 647  
Db 3370 ACAGTGTTCCTAGATGGCAAGAGTAACCACTAATGTCCAGTGGCTCCGAGCACAC 3429  
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667  
Db 3430 CTAGGCATTTGTCTCCCAAGAGCCCATCTGTTTGTAGCTGCAGCATCGCTGAGAATCGT 3489  
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687  
Db 3490 TATGAGACAAACAGCGGGTCTGTCTCCAGGATGAGATCGAGAGGCGCGGAAGAACCC 3549  
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysThrAsnThrGlnValGlyLeuLys 707  
Db 3550 AACATCCACCACTTCATTGAGTCACTGCTGCAAAATACACACAGAGTAGAGACAAA 3609  
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727  
Db 3610 GGGACTCAGCTCTGGGTGGGCGAGAAAACAACGATCGCATAGCAGCGCCCTCGTCAGA 3669  
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747  
Db 3670 CAGCCTCATTTTCTCTCTGATGAAGCAACATCAGCGCTGATACAGAGAGTGAAG 3729  
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
Db 3730 GTTGTCCAGAGCTCTGGCAAGCCAGAGAGGCGGACCTGCTGATGTGATCGCTCAC 3789  
Qy 768 ArgLeuSerAlaIleGlnAsnAlaPheLeuIleValValLeuHisAsnGlyLysIleLys 787  
Db 3790 CGCCTGTCCACCATCCAGAACGACAGCTTGTATCGTGTGATTCAGATGGCAAGGTCAAG 3849  
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleThrPheLysLeuValAsn 807  
Db 3850 GAGCATGGCACCACCGAGCAGCTGTGGCACAAAGAGCATCTATTCTTCAATGTCAGT 3909  
Qy 808 AlaGlnSer 810  
Db 3910 GTGAGGCT 3918  
RESULT 10  
CRUPGP1165 4304 bp mRNA linear ROD 27-APR-1993  
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.  
DEFINITION M59253  
ACCESSION M59253.1 GI:191154  
VERSION multidrug resistance glycoprotein; p-glycoprotein.  
KEYWORDS Chinese hamster, cDNA to mRNA, clone ADX165.  
SOURCE Cricetus sp.  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;

REFERENCE 1 (bases 1 to 4304)  
 AUTHORS Devine,S.E., Hussein,A., Davide,J.P. and Melera,P.W.  
 TITLE Full length and alternatively spliced pgpi transcripts in multidrug-resistant Chinese hamster lung cells  
 J. Biol. Chem. 266 (7), 4545-4555 (1991)  
 MEDLINE 91154285  
 PUBMED 1671863

FEATURES  
 source Location/Qualifiers  
 1..4304  
 /organism="Cricetulus sp."  
 /db\_xref="taxon:10031"  
 /cell\_type="fibroblast"  
 /tissue\_type="lung"  
 1..4304  
 /gene="pgp-1"  
 111..3941  
 /gene="pgp-1"  
 /codon\_start=1  
 /product="p-glycoprotein"  
 /protein\_id="AAA37004.1"  
 /db\_xref="GI:191155"  
 /translation="MEPEDEFGRKDKNFKMGRKSKKKEKPKPVVSVFTMFRVAGWLDRLVMTGLTAAIHGHVALPLMLVFGDMTDSFASVGNIPNTNATNNAATVNASDIFGKLEEMTYAYTYTGAGVLIVATIQVSFWCLAGRQIHKIQKPFHAINWQELGWFVDHVGELNRLTDDVSKINEGIDKIQMFFQAMATFGGFIIGFTGWMKLTLLVLAISPVGLSAGIAWAKILLSFTDKELQAYKAGAVAEVLAAIRTVAFGGQKELERYNNNLEEKARIGIKKAITANTSMGAAPLLIVASVALAFWYGTSLVISEYIGIOVLTVFFA VLIAPFSIQASPNTEAFANARGAAVEIFNII DNKPSIDSFKNQYKPDNIKNLEPKNIHFSYPRKDVOLKGLNLKQVSGQTVALVNSCGKSTTVQLLQRLYLDTEGVVSI DQDITIRNRYLRIVGVSVQEPVLFATTAENIRYKRENTVMDIEKAKEANAYD FIMKLPHKFDLVGRGQAGQGGQRIATARALVNPXKILLDEATSDALDTESEAV QAALDKAREGRTTIVIAHRLSTVRNADIAGFDGGVIVEQGNHEELMEKIGIYFKLVMTOTAGNETELGNEVECKNEIDNLSKSDASSLIRRRSTRSRIRGPHDODRKLSTK EALDEDVPIIFSWRLIKLNSSEWPFVVGIFCAIUNGALQPAFSIIKSVKVGVRNT DDETKRHSNFSLEFLILGVISFIIFLQGFTFGKAGEIITKRLRYVWFKSMRLRQDV SWFNPNKNTGALTIRLANDAGQVKGTGARLAVITQNIAMLTGIIISLIYGWOLTL LLLAIVPTIAIAGVVMKMLSQALKDKKELEGSKIATEAIEINFRTVSVLTREQKFE NMYAQSLQIPYRNALKKARHVFITPSFTQMMYFSYAACFRFGALVARELKMFTENVL LVFSAIIVGMAAGVGVSPADYAKAKVSASHIMIIEKVPSIDSYSGVQLKPNLTLEG NVKNEVVPNTPTDIIPIVLOGLNLEVKKGOTLALVSGCGCKSTVQLLRFERFDPA GTVFDJGGEVQLNQLWMLRHIGIVSQPIFLDCSIAENIAYGNRSVVSQDEIERAA KEANTHQITESLPDKYNTRVGDKGTQLSGGQKRIATARALVROPHKILLDEATSDALD TESERKVOEALDKAREGRTTIVIAHRLSTVRNADIIVVQNGKVEHGHQQLLAQKG IYFSMVSVQAGAKR"

BASE COUNT 1275 a 895 c 1031 g 1103 t  
 ORIGIN Map position 1q26.

Alignment Scores:  
 Pred. No.: 5,076-176 Length: 4304  
 Score: 2476.00 Matches: 468  
 Percent Similarity: 77.52% Conservative: 170  
 Best Local Similarity: 56.87% Mismatches: 171  
 Query Match: 60.70% Indels: 14  
 DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x CRUPGP1165 (1-4304)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1458 ATCGACGGACAGACATACAGGACCAATGTGAGTATCTCGCGGAAATCATCTGGGGTC 1517

Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyAsp 41  
 Db 1518 GTGAGTCAGAACCTGTGTGTTGGCACACATACGCTCGTGAAGAACATTCGCTATGGCCGA 1577

Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 Db 1578 GAAATGTCTACTATGACGAGATCGAGAAAGCTGTCAAGGAAGCAATGCCTATGACTTC 1637

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1638 ATCATGAAACTGCCCATAAATTTGACACTCTTTGTTGGTGGAGAGGACACAGCTGAGT 1697

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1698 GGAGGACAGAAACACAGAAATGCCATTGCTGTCGCTCGCCCAACCCCAAGATCCTT 1757

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 1758 TTGTTGGATGAGGCGACATCAGCCTTGACAGAAAGTGAAGCCGTGGTTCAGGCTGT 1817

Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArgLeuSerThrIle 141  
 Db 1818 CTGGATAGGCTTAGAGAGGCGGACTACCATTTGTATAGCCACCGCTTGTCTACAGTT 1877

Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
 Db 1878 CGAAATGCTGACATCATAGCTGGTGGTTCATTTGTGAGCAGCAAGAAATCAT 1937

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
 Db 1938 GAAGAGCTCATGAGAGAGAGGCGCATTTTCAAACTTGTCTATGACACAGCAGCAGGA 1997

Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 191  
 Db 1998 AATGAATGTAATAGGAATGAAGTTCGTGAGTCTAATAAATGAATTCATAATTAGAC 2057

Qy 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209  
 Db 2058 ATGTCTCTCAAAAGATTTCAGCATCCAGTCTAATCAGAAAGAGATCAACTCGCAGAAATATC 2117

Qy 210 LysSerAppPheIleAspLysAlaGluGluSerThr--GlnSerLysGluLeuSerLeu 228  
 Db 2118 COTGGACCATCATGACACAGAGGAGGCTTAGTCAAAAGAGCGCTTGGATGAAGATGTA 2177

Qy 229 ProGluValSerLeuLysIleLeuLysLysLeuAsnLysProGluTyrProPheValVal 248  
 Db 2178 CTTCAATTTCTTTTGGCGTATCCTAAGATTGAATTCATCTGAATGGCCATATTTTGG 2237

Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
 Db 2238 GTTGGTATATTCTGTGCCATAGTAATAAGGAGCTTGCAACACGACATCTCTCAATAATTT 2297

Qy 269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287  
 Db 2298 TCTAAGTTGTAGGGGTTTTCACAGAAATACTGATGATGAACCAACCAACGACATGATAC 2357

Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
 Db 2358 AACTTGTCTTCTTATGTTTCTGATCTTGGGGTCATCTTTTATACATTTTTTCTT 2417

Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
 Db 2418 CAGGCTTCACATTTGGCAAAAGCTGGAGAGATCCTCACCAGGAGCTCGGATACATGTT 2477

Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
 Db 2478 TTCAATCCATGCTGAGACAGATGTGAGTGTGTTGATAACCCCTAAACCAACCACTGGA 2537

Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValaThrGlySerArg 367  
 Db 2538 GCATTGACACACAGGCTGCCAACCATGCTGCTCAAGTTAAAGGGGCTACAGAGCCAGA 2597

Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
 Db 2598 CTTGCTGTCTATACCCAGAAATATACCAATCTTGGGACAGGAATCATCATATCCCTAATC 2657

Qy 388 TyrGlyTyrPheLeuMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
 Db 2658 TATGCTGTCAGTTCACACTTCTACTCTTAGCAATTTGTCCCATCATTCATTCAATACAGGA 2717

Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
 Db 2718 GTGGTTGATGATGAAATGTTGCTGGCAAGCAGCACTAAAGATGAAGAGGAGCTAGAGGT 2777

```
QY 428 AlaGlyIysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 2778 TCTGGGAAGATTCTACTGAAGCAATAGAGAACTTCGCAGTCTCTCTTTGACTCG 2837
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2838 GAGCAGAAGTTTGAAATATGTATGTCAGAGCTTGAGATACCATACAGAAATGCTCTG 2897
QY 468 LysLysAlaGlnIleLeuGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2898 AAGAAAGCACACGCTTTGGGATCACATTCTCTCCAGCCAGGCCATGATGATTTTTCG 2957
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2958 TATGCTGCTTTGTTTCGGTTTGTGCTTACTTGTGGCAGCGGAATTAATGACATTGAA 3017
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 3018 AATGTTCTATTAGTATCTCAGCTATTGCTTTGGTGCCATGGCAGTGGGCGAGTCAGT 3077
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3078 TCATTTGCTCTGATGCTGCANAGCAAGTCTCGGCATCGCATATCATCATGATCAIT 3137
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3138 GAAAGAGTCCCTTCCATTCAGCTACAGCACGCGAGGCTGAAGCCTAATACATTGGAA 3197
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3198 GGAATGTGAATTTAATGAAGTCGTGTTCACTACCCAGCCGCTGACATCCAGTGG 3257
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3258 CTTTCAGGGCTGAACCTGGAGGTGAAGAGGCGCAGACGCTGCGTGGCGCAGCAGT 3317
QY 608 GlyCysGlyLysSerThrSerValGlnLeuGlnLeuArgLeuTyrAspProValGlnGly 627
Db 3318 GGCTGGGGAGAGCAGACGTGTCAGCTCTCGAGCGGTCTTACGACCCCATGCGCGGA 3377
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3378 ACAGTGTCTTAGTGGCAAGAAAGTAACCAACTAATGTCTCAGTGGCTCCGAGCAGCAC 3437
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3438 CTAGGCAATTGTGCCAAGAGCCCTCTCTTTGACTGCGACATCCTCGAGAATCGCT 3497
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAlaAla 687
Db 3498 TATGGAGACAAACAGCGGGTCTGTCTCCAGGATGAGATCGAGAGGCGCGCAAGGAAGC 3557
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3558 AACATCCACAGTTCATTAGTCACTGCTCGCAAAATACAACACAGAGTAGGAGACAAA 3617
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3618 GGGACTCAGCTGTGGGTGGCGAAGAACACGATCGCCATAGCAGCGCCCTCGTCAGA 3677
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3678 CAGCCTCACATTTTGTCTCTGGATGAAGCAACATCAGCCCTCGGATACAGAGATGAAG 3737
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3738 GTTGTCCAGGAAGCTCTGGCAAAAGCCAGAGAGCGCACCTGTCATGTGATCGCTCAC 3797
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3798 CGCCTGTCCACCATCCAGAACGAGACTTGTGTTGATTCAGATGCGCAGGTCAAG 3857
QY 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
```

```
Db 3858 GAGCATGGCACCACGAGCTGCTGGCACAGAAAGCATCTATTTTCCATGGTCACT 3917
QY 808 AlaGlnSer 810
Db 3918 GTGCAGGCT 3926
RESULT 11
OAU78609 3858 bp mRNA linear MAM 11-OCT-2001
LOCUS Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
DEFINITION
ACCESSION U78609 GI:2149086
VERSION U78609.1
KEYWORDS
SOURCE Ovis aries.
ORGANISM Ovis aries.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
REFERENCE 1 (bases 1 to 3858)
AUTHORS Longley, M., Phua, S.H., van Stijn, T.C. and Crawford, A.M.
TITLE Isolation and mapping of the first ruminant multidrug resistance
Genes
JOURNAL Anim. Genet. 30 (3), 207-210 (1999)
MEDLINE 99371931
PUBMED 10442984
REFERENCE 2 (bases 1 to 3858)
AUTHORS Longley, M. and Crawford, A.M.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1996) Biochemistry Department, University of
Otago, Dunedin, New Zealand
FEATURES
Location/Qualifiers
1..3858
/organism="Ovis aries"
/db_xref="taxon:9940"
1..3858
/gene="mdr1"
1..3858
/gene="mdr1"
/codon_start=1
/product="multidrug resistance protein-1"
/protein_id="AAB58489.1"
/db_xref="GI:2149087"
/translat ion="MDLEGDRNGRAAGNFKLRDKRPFKSKDEKKEKRPVSTFTMF
RYSNWLRLCMVLGTAAI IHGAGLPLMLVFGDMTDSFAGAGNLGNITLNSINTST
IDREYCKLEKEMTVAYYVGIGAGVLIAYIQVSPWCLAAQROVHRIKQPHAI
MQQSIGPDVHDVKLNRLNDVSKNEGIDKIMFFQMAVTFFTGFIQFTGNW
LTLVILALIPVLGSAAIWAKILSSFTDDKULAYAKAGAAVEVLAALKTIVAFGGQK
KELERYNNLEAKRIGIKAITANISMAAFLLIYASVALFYGTSLYSREYSIG
QVLTVFSLVIGAFSIGQASPNI EAFANARGAAVEFKI IDNKPSIDSYNTGHPDN
IKGNLEPRNVHFHYPSRNEVKILKGLNVKVGSGOTVALVNSGCGKSTTVOLMORLYD
PTEGMVSDIGDITINRVYIREIGVVSQSPVLPATIAENIRYGRDVTMDELOKA
VKEANAYDFIMKLPNKFDLTVGERQAQSGQKQRIARALVRNPKILLDEATSAI
GTESEVQAAALDKARKGRTTIVIAHRLSTVNRNADIAGLDGVIVEEGSHDLMGKR
GYFKLVTMQTKNELENTFVLSKTDLYMSQDSRSLRRKTRRSIRGSQS
QDKLSTEETLDESVPVPSFWRLKLNITEMPYFVGVFCAI INGALOPAFSVFSRI
IGITRNNDDETKQNSLRFLLIGIISFITFFLOGFTFGKAGETLTRRLRYLVF
RSMRLQDVSFEDDPKNTGALTTLANDAAQKAVGSLAVITQNIANLTGIIISL
IYQWETULLLAIPIITAVAGVIEKMSLQGLKDKKELEGAKIATEIENFRVVS
LTRERFEYMQALQVYPYRNSLRRAHVFGITFTIQMAYFSVAGCFRFGALVLAQG
IMEFQDVLIVFSAVVFGAMVQVSTFADYAKAKVSAHVINIIEKTLPLDSVTEG
LKPSTVEGSVAFNDVFNYPTRPDVPLVRLGLSLVKVKGOTLALVSGSCGKSTVQLL
ERFYDPLAGTVFIDGKEVKQLNVOLRAHMGIVSQEPLFDCSITENAYGDNRSVVS
OEETEHAKENAIHSFLEMLPKYNTRVGDGTOLSGOKORIAARALVRQPHILL
DEATSAIDTESEKVVQEAALDKARERTCIVIAHRLSTIQNALDLIVVFQNGRIKEGTH
COLLAQKGIYFTWVSQAGTKRQ"
BASE COUNT 1072 a 853 c 1002 g 931 t
ORIGIN
Alignment Scores: 1.74e-175 Length: 3858
Pred. No.: 2468.00 Matches: 475
Score:
```



Percent Similarity: 76.73% Conservative: 158  
Best Local Similarity: 57.58% Mismatches: 174  
Query Match: 60.51% Indels: 18  
DB: 4 Gaps: 4

US-09-873-409-2 (1-812) x OAU78609 (1-3858)

QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyArgAspHisIleGlyVal 21  
DB 1372 ATCGACGGACAGATATCAGGACCACTCAACGTAAGGTATCTCGGAGCATATTCGGGGTG 1431  
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleIysTyArg 41  
DB 1432 GTGAGTCAGAGCCCTGTGCTGTTGCCACACGATAGCTGAGAACATTCGTCAGCGCCT 1491  
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyArgPhe 61  
DB 1492 GAAGATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551  
QY 62 IleMetGluPheProAsnIysPheAsnThrLeuValGlyGlyGlyAlaGlnMetSer 81  
DB 1552 ATCATGAACTGCCTAATAAGTTTGACACCTGTTGGAGAGAGAGAGAGAGAGAGAGAG 1611  
QY 82 GlyGlyGlnIysGlnArgIleAlaAlaArgAlaLeuValArgAsnProIysIleLeu 101  
DB 1612 GTGGACAGACAAAGAAATGCCATTTGCCCGGCCCTGTTGCCAACCCCAAGATCTCTC 1671  
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
DB 1672 CTGTTGGATGAGCCACCTCAGCGTGGCCACGAGAGTGAAGCGGTTCAGCGGCC 1731  
QY 122 LeuGluYsAlaSerIysGlyArgThrIleValValAlaHisArgLeuSerThrIle 141  
DB 1732 CTGGAAGCCAGAAAGCCGACCATTTGTATAGCGCATCGTCTGTCCACGGTT 1791  
QY 142 ArgSerAlaAspIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
DB 1792 CCAATGCTGCATCTCTGCTGTATGATGATGATGATGATGATGATGATGATGATGAT 1851  
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyThrSerLeuValMetSerGlnAspIleLys 181  
DB 1852 GATGAGCTCATGGGAAAGAGCATTACTTCAACTTGTACAACTGTACAACTGCAG--ACAAA 1908  
QY 182 LysAlaAspGluGlnMetGluSerMetThrTyThrSerThrGluArgLysThrAsnSerLeu 201  
DB 1909 CGAAATGAACCTGAATTGGAAATATCCCTGTTGATCCCTTAAGTAAATCTGATGACTTG 1968  
QY 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214  
DB 1969 TACATGCTCTCCCAAGATTCA---AGATCCAGTCTAATCAGAAAGAAATCAACTCGCAGG 2025  
QY 215 -----ApLysAlaGluGluSerThrGlnSerLysGluIle 226  
DB 2026 AGTATCCGGATGCCAAAGCCAGGACAGAAAGCTCAGTACAGAAAGAACTTTGGATGAA 2085  
QY 227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246  
DB 2086 AGTGATCTCCCGTTCTTTTCGAGGATCTGAAGCTGAATATAACTGAATGGCCTTAT 2145  
QY 247 ValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIle 266  
DB 2146 TTTGTGTTGGTGTATTTTGGCGCCATTATAAATGGAGCTCTGCAACCCAGCATTTCTCAGTC 2205  
QY 267 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys---ThrThrLeuLysHis 285  
DB 2206 ATATTTTCAAGGATATAGGAATCTTCCACAGAAATGACAATGACGAAACCAACAGCAG 2265  
QY 286 AspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheValSerTy 305  
DB 2266 AACAGTAACCTGTTTCACTGTGTTTCTAATCCTTGGAAATATTTCTTTTATCATTT 2325  
QY 306 PheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325  
DB 3406 ATCGCTACGGAGACAAACAGCGCGGTGCTGTCCAGGAGGAGATTGAGCACGCGCCAAAG 3465

DB 2326 TTCCTTCAGGGCTTCACGTTTCGGGAAACAGCGGAGAGATCTCTCACAGGGCGGTGCGGTAC 2385  
QY 326 LeuAlaPheLysAlaMetLeuTyArgGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345  
DB 2386 CTGCTTTTCAGTCGATCTGAGACAGCATGTGAGCTGGTTCGATGATGACCTTAAACACACC 2445  
QY 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365  
DB 2446 ACGGAGCATTTGACACACAGGCTTCCCAATGACGCGCTCAAGTTTAAAGGGGCTGTAGGT 2505  
QY 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385  
DB 2506 TCCAGCTTCGTGTAATTTACCCAGAAATATTGCAAAATCTTGGACAGGCAATTAATATCC 2565  
QY 386 PheIleTyArgGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405  
DB 2566 CTCATCTATGCTGCGCAGTTAACACTTCTCTGTTAGCAATTTGACCATCATTCAGTA 2625  
QY 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425  
DB 2626 GCAGGAGTTATTGAAATGAAAATGCTGTCTGGACAAAGCCCTGAAAGATAAGAAAGACTG 2685  
QY 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445  
DB 2686 GAAGTGTCTGGAAATTTGCCACTGAACAAATAGAAAATCTTCGGACCGGTTGTTCTTTG 2745  
QY 446 ThrArgGluLysAlaPheGluGlnMetTyArgGluGluMetLeuGlnThrGlnHisArgAsn 465  
DB 2746 ACTCGGAGGAGAGGTTTGAATATATATATGATGATGATGATGATGATGATGATGATGAT 2805  
QY 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyArgAlaPheSerHisAlaPheIleTy 485  
DB 2806 TCCTTGAGGAGAGCGCATGCTCTTTCGAAATTTACATTTTCCATCAGCAATGATGAT 2865  
QY 486 PheAlaTyArgAlaGlyPheArgPheGlyAlaTyLeuIleGlnAlaGlyArgMetThr 505  
DB 2866 TTCTCTATGCTGGCTGTTCCGGTTTGTGCTACTTGTGTCGCTGTCGTCGTCGTCGTCGTC 2925  
QY 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyArgGlyAlaMetAlaIleGlyLys 525  
DB 2926 TTTTCAGGATGTTCTCTAGTGTCTCAGCTGTGTCTTGTGTCGTCGTCGTCGTCGTCGTC 2985  
QY 526 ThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545  
DB 2986 GTCACTACATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3045  
QY 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565  
DB 3046 ATCATTTGAAAAATTTCTCTGATTCACAGCTACAGCGGAAAGGCTTAAAGCCGAGTACA 3105  
QY 566 CysGluGlyAsnLeuGluPheArgGluValSerPheTyProCysArgProAspVal 585  
DB 3106 GTGGAAAGAGTGTGGCATTTAATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3165  
QY 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605  
DB 3166 CCAGTGCTTCGGGGCTGAGCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3225  
QY 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgProVal 625  
DB 3226 AGTAGTGCTGTGGAGAGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3285  
QY 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645  
DB 3286 GCTGGCAGAGTGTATTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3345  
QY 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665  
DB 3346 GCACACATGGGCATCGTCTCCAGGAGGCCATTTCTGTTGACTGAGCATCGCGAGAAT 3405  
QY 666 IleAlaTyArgLysAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685  
DB 3406 ATCGCTACGGAGACAAACAGCGCGGTGCTGTCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 3465





QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSerGln----- 178  
Db 2027 AGTGAGCTGATGAAGAAGGAGGATCTCTTACAGATCGACAGCAGCA 2086  
QY 179 -----AspIleLysAlaAspGluGlnMetGluSerMet 190  
Db 2087 AGCCAGATCCTGTCAGAGAAATTTGAAGTTGAGCTTAAGTGACAAAGGCTCTGGAGAT 2146  
QY 191 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 210  
Db 2147 GTGGCCCCAAATGGCTGGAAGACGACCATATTTAGGAATTTCTACAAAGAAAGTCTTAA 2206  
QY 211 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluLysSerLeuProGlu 230  
Db 2207 AGTCCACATCAGATAGGCTGATGAAGAAACCAATGAATCTTATGCAAAACGTCACCA 2266  
QY 231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGly 250  
Db 2267 GTGTCTTTTCTGAAGGTCTTAAACTGAATATAAACAGAGTGGCCCTACTTTGTGTGGGA 2326  
QY 251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270  
Db 2327 ACAGTCTGTGCCATTCCTATCGAGCCCTCCAGCCGCTTCTCCATCATCTGTCTGAG 2386  
QY 271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290  
Db 2387 ATGATAGCTATCTTGGCCCTGGGATGACGCGATGAAGCAAAAGTGTAAATGTTTC 2446  
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310  
Db 2447 TCCTCGTCTTCTTGGCCCTAGAGTCCCTCTCTTACTTCTTCTTCTTCTTCTTCTTCTTCT 2506  
QY 311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330  
Db 2507 ACGTTTGGGAAGCTGGAGAGATCTTACCACAGAGCTCGGTCCATGCGCTTAAAGGG 2566  
QY 331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350  
Db 2567 ATGCTAAGGCAGACATGAGCTGGTTTGTATGATCATATAAAACAGTACTGGAGCACTTCT 2626  
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370  
Db 2627 ACAAGACTCGCCACAGATGCTGCGCAAGTCCAAAGAGCCACGGAACCAAGTTGGCTTTA 2686  
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 390  
Db 2687 ATTGACAGACACAGCCACCTTGGACCGGTATTTATATATATATATATATATATATATAT 2746  
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410  
Db 2747 CAACCTGACACTTCTGCTGTATCGGTGTGTTCCATTCATTCCTAGCAGGAATTTGTAA 2806  
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430  
Db 2807 ATGAAAATGTTGGCTGCAATGCAAGAGAGATATAAAGGAAATGGAAGCTGCTGGAAG 2866  
QY 431 IleAlaThrGluAlaLeuGluAsnIleAtrThrIleValSerLeuThrArgGlyLysAla 450  
Db 2867 ATTGCAACAGAGCAATAGAAAATATTCGAACTGTGTATCTTGTACCCCAAGAAAGAAA 2926  
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470  
Db 2927 TTTGAGTCAATGATGTTGAAAATTTGATGGACCTTACAGGAATTCGGTGGCGGAAGGCA 2986  
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490  
Db 2987 CACATCTACGCCATCACTTTTAGCATCTCCCAAGCATTCATGATTTTCTTATGCTGGC 3046  
QY 491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510  
Db 3047 TGTTTTCGATTTGCTTCTTACCTAAATGTGAATGGACATATCGCTTCAAAAGATGTCATT 3106  
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530

Db 3107 CTGGTCTTTTCTGCAATTTGCTTGGCCGGTGGCTCTAGGACACGCCAGCTCATTTGCT 3166  
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550  
Db 3167 CCGGACTATGCAAAAGCCCAAGCTGCTGCAGCATATTTGTTTTCAGCTGTTTGAAGACAA 3226  
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570  
Db 3227 CCTCTGATTGACAGCTACAGTGGAGAGGGCTGTGGCCCTGATAAGTTTGAAGGAGCCGTG 3286  
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590  
Db 3287 ACATTTAATGAAGTCGTGTTCAACTATCCACCGGGCCAGCTGCCAGTGCCTCAGGGG 3346  
QY 591 LeuSerLeuSerIleGluArgGlyThrValAlaPheValGlySerSerGlyCysGly 610  
Db 3347 CTGAGCCTTGAAGTGAAGAGGCCAGAGCTGGCCCTGGTGGGCGAGCAGTGGCTCGGG 3406  
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630  
Db 3407 AAGACACAGTGGTCCAGCTGCTCGAGGGCTTCTATGACCCCATGGCTGGATCAGTGTCT 3466  
QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650  
Db 3467 TTGATGTTGCAAGAGCAAGAACTCAATGTCCAGTGGCTCCGAGCTCACTGGGCAAT 3526  
QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670  
Db 3527 GTGTCCAGGAACCATCTCTTTTACCTGCAGCATCCGAGAAACATCGCTATGGAGAC 3586  
QY 671 AsnSerArgValProLeuAspGluLysGluAlaAlaAsnAlaAlaAsnIleHis 690  
Db 3587 AACAGCCGGTCTGCTCATGATGAGATTGTGAGGCGAGCCCAAGGAGCCCAACATCCAC 3646  
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710  
Db 3647 CCTCTCATCGAGACCTGCCCAAAATATACACAAGAGTAGGAGACAAGGGGACGCGAG 3706  
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuLysProLys 730  
Db 3707 CTCTTGGGGCCAGAGCAGAGGATTCCTATCGCCGAGCCCTCATCAGACAGCTCGG 3766  
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750  
Db 3767 GTCTACTGCTGGATGAAGCCACGTCAGCTCTGGATCTAGAGTGAAGGTTGTCCAG 3826  
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770  
Db 3827 GAAGCACTGGACAAAGCCAGGGAAGCCGACCTGATTTGTATGCTCACCCTGTCTCC 3886  
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790  
Db 3887 ACATCCAGAACCGGACTTGTATGTTGATGTAGACGGCAAGGTCAAGGACGACGGC 3946  
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810  
Db 3947 ACCCACCAGCAGCTGTGGCGCAGAAAGGCACTATTTCTCAATGTCTCAACATCCAGGCC 4006  
RESULT 13  
LOCUS AX024455 3924 bp DNA linear PAT 15-SEP-2000  
DEFINITION Sequence 2 from Patent DE19857750.  
ACCESSION AX024455  
VERSION AX024455.1 GI:10184623  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 3924)  
JOURNAL Streimel, W.  
Patent: DE 19857750-A 2 24-FEB-2000;

STREMMEL WOLFGANG (DE)		Location/Qualifiers	
source	1..3924	/organism="Homo sapiens"	
		/db xref="taxon:9606"	
		/note="MDR3/menschliche MDR-Genfamilie"	
BASE COUNT	1145 a	790 c	977 g 1012 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.74e-175	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	6	Gaps:	3
US-09-873-409-2 (1-812) x AX024455 (1-3924)			
Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21
Db	1395	ATTGATGGCCAGGATATTAGGAATCTTAATGTAACTATCTGAGCGAAATCATTTGGTGTG	1454
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1455	GTGAGTCAGGAGCGGTGCTGTTTCCACCACAAATTCGTGAAATAATTTGTTATGGCCGT	1514
Qy	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
Db	1515	GGAAATGTAACTGATGATGATGAAGAAGCTGTCAAGAGGCCAACGCCCTATGAGTTT	1574
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1575	ATCATGAATTTACACAGAAATTTGACACCCTGTTGGAGAGAGGGGCCAGCTGAGT	1634
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1635	GGTGGCAGAGCAGAGATCGCCATTCACGTGCGCTTGGCAACCCCAAGATCCCTT	1694
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1695	CTGCTGGATGAGGCCAGCTCAGCATTTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT	1754
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1755	CTGGATAAGCCAGAGAGGCCGACCATTTGTGATAGCACACCCAGCTGTCTACGGTC	1814
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1815	CGAAATGCAGATGTCATCGCTGGGTTTGGAGTGGAGTAATTTGTGGAGCAAGAGCCAC	1874
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	175
Db	1875	AGCGAACTGATGAAGAGGAGGGGTGACTTCAAACTGTGCAACATGACGACATCAGGA	1934
Qy	176	-----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr	192
Db	1935	AGCCAGATCCAGTCAGAAATTTGAACATAATGATGAAGGCTGCCACTAGAATGCC	1994
Qy	193	SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser	211
Db	1995	CCAAATGGCTGGAAATCTCGCCTATTAGGCATTTCTACTCAGAAACACCTTAAAAATTC	2054
Qy	212	AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu	230
Db	2055	CAAAATGTGTGAGAGAGCCCTTGATGTGGAAACCGATGGACTTGAAGCAATGTGCCACCA	2114
Qy	231	ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly	250
Db	2115	GTGTCCTTTTGAAGGTCCTGAACTGAAATAAAACAGATGGCCCTACTTTGTCGGGA	2174
Qy	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	3255	AAAGACACGGTGGTCCAGCTCTCGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCCT	3314

QY	631	PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle	650
Db	3315	CTCGATGGTCAAGAAGCAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC	3374
QY	651	ValProGlnProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp	670
Db	3375	GTCTCTCAGGAGCGTATCTCTATTGTGATCGACATGTCGCGAGAATATTGGCTATTGGAGAC	3434
QY	671	AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis	690
Db	3435	AAcAGCGGGTGTATCACAGATGMAATTGTGAGTCAGCGCAAGACTGCCCAATACAT	3494
QY	691	SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln	710
Db	3495	CCTTTTCATCGAGACGTTACCCACAAATATGAACAAGAGTGGGAGATAAGGGACTCAG	3554
QY	711	LeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys	730
Db	3555	CTCTCAGGAGGTCAAAACAGAGGATTGCTATTGTCCCGAGCCCTCATCAGACAACCTCAA	3614
QY	731	IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln	750
Db	3615	ATCTCTCTGTTGATGAGCTACATCAGCTCTGGATCTGGAATCTGAAAGTGAAGAGTTGTCCAA	3674
QY	751	HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer	770
Db	3675	GAAGCCCTGGACAAAGCCAGAGAAGCCGCACCTGCTATTGTGATTGCTCAGCGCTGTGCC	3734
QY	771	AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleGlyGluGlnGly	790
Db	3735	ACCATCCAGAATCCAGACTTAATAGTGGTGTTCAGAATGGGAGATCAAGGAGCATGGC	3794
QY	791	ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	810
Db	3795	ACGATCAGCAGCTGCTGGCACAGAAGGCATCTATTTTCAATGTCAGTCTCAGGCT	3854
RESULT 14			
AX409652	AX409652	3924 bp	DNA
LOCUS	Sequence 2299	from Patent WO0229103.	
DEFINITION	AX409652		
ACCESSION	AX409652.1	GI:21442357	
VERSION			
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1		
JOURNAL	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.		
FEATURES	Gene expression profiles in liver cancer		
source	Patent: WO 0229103-A 2299 11-APR-2002;		
	GENE LOGIC INC (US)		
	Location/Qualifiers		
	1..3924		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="EMBL/GenBank Accession No. M23234"		
BASE COUNT	1145 a	790 c	977 g
ORIGIN	1012 t		
Alignment Scores:			
Pred. No.:	2,74e-175	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	6	Gaps:	3
US-09-873-409-2 (1-812) x AX409652 (1-3924)			
QY	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21

Db	1395	ATTGATGGCAGAGATATTAGGAACCTTAAATGTAACATATCTCAGGGAATCAATCATTTGGTGTG	145
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleuYerGlyArg	41
Db	1455	GTGAGTCAGGAGCGGTGTGTTTTCCACCACAAATGTCGAAATATTTTGTATGGCCGT	1514
Qy	42	AspAspValThrAspGluMetGluArgAlaAaArgGluAaAsnAlaTyrAspPhe	61
Db	1515	GGAAATGTAAACCATGATGAGATAAAGAAGCTGTCAAGAGAGGCCACGCCCTATGAGFTT	1574
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1575	ATCATGAAATACACACAGAAATTTGCACCCCTGGTTGGAGAGAGAGGGCCAGCTGAGT	1634
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1635	GGTGGGCAGAAAGCAGAGGATCGCATTCGACGTGCTGCTGTTGCGAACCCCAAGATCCTT	1694
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1695	CTGCTGGATCAGGCCACGTGACGATTTGGACACAGAAAGTGAAGCTGAGGTACAGCGAGCT	1754
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1755	CTGGATGAAGCCACAGAAAGCCGGACACCATTTGTATGATACACACCGACTGTCTACGGTC	1814
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1815	CGAAATGCAGATGTCATCGCTGGTGGTTGAGGATGAGTAAATTTGTGGACAGGAAGCAC	1874
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	175
Db	1875	AGCGAATCGATGATGAAGAAGGAGGAGGTTGACTTTCAAACTTGTCAACATCGACACATCAGGA	1934
Qy	176	-----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr	192
Db	1935	AGCCAGATCCAGTCAGAGAAGATTTGAACATAATGATGAAGAAGCTGCCACTAGAAATGCC	1994
Qy	193	SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser	211
Db	1995	CCAAATGGCTGGAAATCTCGCTATTTAGGCATTTCTACTCAGAAAAACCTTAAAAAATTC	2054
Qy	212	AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu	230
Db	2055	CAAAATGTCTCAGAAAGACCTTGATGTGGAACCCGATGGACTTGAAGCAAAATGTGCCACCA	2114
Qy	231	ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly	250
Db	2115	GTGTCTCTTCTGAAGGTCCTGAACATCGAATAAACAAGATGCCCTACTTTGTCTGGGGA	2174
Qy	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	2175	ACAGTATGCCATTCGAATGGGGCTTCAGCGGCATTTTCAGTCATATTTCTCAG	2234
Qy	271	IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr	290
Db	2235	ATCATAGCGATTTTGGACCGCGCATGATGTCAGTGAAGCAGCAAGTGCACATATTC	2294
Qy	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu	310
Db	2295	TCCTTGATTTCTTATTTCTGGGAATATTTCTTTTTTTTACTTTCTCTCTCAGGGTTTC	2354
Qy	311	PheTyrGlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla	330
Db	2355	ACGTTTGGGAAAGCTGGCGAGATCCTCACAGAGAGACTCGGTCAATGGCTTTTAAAGCA	2414
Qy	331	MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr	350
Db	2415	ATGCTAAGACAGGACATGAGCTGGTTGATGACCATATAAACAAGTACTGGTGCATCTTCT	2474
Qy	351	ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
Db	2475	ACAAGACTGCCACAGATGCTGCCAAGTCTCAAGGAGCCACAGGAACCAAGGTTGCTTTA	2534



Pred. No.:	2.74e-175	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	9	Gaps:	3

US-09-873-409-2 (1-812) x HUMMDR3 (1-3924)

Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21
Db	1395	ATTGATGGCAGGATATTAGGAACCTTTATGTAACATATCTCAGGGAATCATTTGGTGTG	1454
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1455	GTGAGTCAGGAGCGGTGCTGTTTCCACCAAAATGCTGGAATAATTTGTTATGGCCGT	1514
Qy	42	AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
Db	1515	GGAAATGTAAACCATGGATGAGATGAAGAAGCTGTCAAGAGGCCCAACGCCCTATGAGTTT	1574
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGluValAlaGlnMetSer	81
Db	1575	ATCATGAATTAACACAGAAATTTGACCCCTGGTTGGAGAGAGGGGCCCGCTGAGT	1634
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1635	GGTGGCAGAACGACGAGGATGCCATTGACCTGCGCTGTTGCGCAACCCCAAGATCCTT	1694
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1695	CTGCTGGATGAGGCCACGTCAGCATTTGGACACAGAAAGTGAAGTCAGGTACAGGCAGCT	1754
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1755	CTGGATAAGCCAGAGAGCGGACCCATTTGTGATAGCACACCGACTGCTCAGGTGTC	1814
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1815	CGAAATGCAGATGTCATCGCTGGTTTGAGGATGGAGTAATTTGTGGAGCAAGGAACCCAC	1874
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal-----	175
Db	1875	AGCGAACTGATGAAGAAGGAAGGGGTGATCTTCAAACTTTGTCAACATGTCAGACATCAGGA	1934
Qy	176	-----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr	192
Db	1935	AGCCAGATCCAGTCAGAGAAATTTGAACCTAAATGATGAAGAGGCTGCCACTAGAAATGGCC	1994
Qy	193	SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser	211
Db	1995	CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAATAATCA	2054
Qy	212	AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu	230
Db	2055	CAAAATGTCAGAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAAATGTGCCACCA	2114
Qy	231	ValSerLeuLeuLysIleLysLeuAsnLysProGluTyrProPheValValLeuGly	250
Db	2115	GTGTCCTTTCTGAAGTCTCTGAACCTGAATAAAACAGAAATGGCCCTACTTTGTGCGGGA	2174
Qy	251	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	2175	ACGATATGTCCTAATGGGGGCTTCAGCCGGCATTTTCATCATATCTCAGAG	2234
Qy	271	IleIleThrMetPheGlyAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr	290
Db	2235	ATCATAGCATTTTGGACCGGCGATGATGTCAGTGAAGCAGCAGCAAGTGAACATATTC	2294
Qy	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu	310
Db	2295	TCTTTGATTTCTTATTTCTGGGAATATTCTCTTTTACTTCTTCTCTCTCTCTCTCTCT	2354

Qy	311	PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheIysAla	330
Db	2355	ACGTTTGGGAAGCTGGCAGATCCTCACCAGAAGACTCGCGTCAATGGCTTTTAAAGCA	2414
Qy	331	MetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThr	350
Db	2415	ATGCTAGACAGGACATGAGCTGGTTGTATGACCATTAACACAGTACTGTGTCATTTCT	2474
Qy	351	ThrLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
Db	2475	ACAAAGCTTGGCAGATGCTGCCAAGTCCAGGAGCCACAGGAACCCAGGTGGCTTTA	2534
Qy	371	LeuThrGlnAsnAlaThrAspMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr	390
Db	2535	ATTGCACAGAATATAGCTAACCTTTGGAACTGGTATTATCATATCATTTATCTACGGTGG	2594
Qy	391	GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu	410
Db	2595	CAGTTAACCTTATTTGCTATTAGCAGTTGTTCCAAATTATTGCTGTGTCAGGAATTTGTA	2654
Qy	411	ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys	430
Db	2655	ATGAAATTTGTTGGCTGGAAATGCCAAAAGAGATAAAAAGAACTGGAAGCTGCTGGAAG	2714
Qy	431	IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla	450
Db	2715	ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTCTTTGACCCAGGAAGAAAA	2774
Qy	451	PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla	470
Db	2775	TTTGATCAATGTAATGTTGAAAATTTGTATGACCTTACAGGAATTTCTGTGAGGAAGCA	2834
Qy	471	GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla	490
Db	2835	CACATCTATGGAATTTACTTTTAGTATCTCACAGCATTTATGATTTTCTTATCCCGGT	2894
Qy	491	GlyPheArgPheGlyAlaTyrIleGlnAlaGlyArgMetThrProGluGlyMetPhe	510
Db	2895	TGTTTTCGATTTGGTGCAATATCTCATTTGTAATGACATATGCGCTTCAGAGATTTATT	2954
Qy	511	IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla	530
Db	2955	CTGGTGTCTTCTGCAATTTGTTGGTCAGTGGCTCTAGGACATGCCAGTTCATTGCT	3014
Qy	531	ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys	550
Db	3015	CCAGACTATGCTAAGCTAAGCTGCTGCAGCCCACTTATTATCATCTGTTTGAAGAGCAA	3074
Qy	551	ProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeu	570
Db	3075	CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCCTGATAAATTTGAAGGAAATATA	3134
Qy	571	GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly	590
Db	3135	ACATTATTAAGTCTGTTCAACTATCCCAACCCAGCAAGCTGCCAGTGTCTTCAGGGG	3194
Qy	591	LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly	610
Db	3195	CTGACCTGGAGGTCAAGAAAGCCAGACACTAGCCCTGGTGGGAGCAGTGGCTGTGGG	3254
Qy	611	LysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu	630
Db	3255	AAAGCAGCGTGGTCCAGCTCTCGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGTCTT	3314
Qy	631	PheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIle	650
Db	3315	CTCGATGTCTAAGAACCAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC	3374
Qy	651	ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp	670
Db	3375	GTGCTCAGGACCTATCTCTATTGACTGTCAGCATTTGGAGAAATATTCCTATGAGAC	3434
Qy	671	AsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis	690

```
Db 3435 AACAGCCGGGTTGTATCAGAGGATGAAATTTGAGTGCAGCCAAAGCTGCCAACATACAT 3494
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db 3495 CCTTTCATCGAGAGCTTACCCCAATATGAAACAAAGAGTGGGAGATAAGGGGACTCAG 3554
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCGAGCCCTCATCAGACAACCTCAA 3614
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db 3615 ATCCTCCTGTGGATGAAGCTACATCAGCTCTGATACTGAAAGTGAAGAGGTGTGCCAA 3674
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db 3675 GAAGCCCTGGACAAAGCCAGAGAGCCGACCTGCTGTTGTTGCTCACCCTGTCTCC 3734
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db 3735 ACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGAAATGGGAGAGTCAAGGAGCATGGC 3794
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db 3795 ACCCATCAGAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 02:09:55  
Job time : 11128.4 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 793.594 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-2

Perfect score: 4079

Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLNRNDIVFKLVNAQSVQ 812

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPRO\_pool/US09873409/runat\_27032003\_115416\_19216/app\_query.fasta\_1.7544  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 4247 @runat\_27032003\_115416\_19216 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3919	96.1	3699	24	Human transporter
2	2465.5	60.4	3924	21	Human ATP binding
3	2465.5	60.4	3924	21	Human MDR-3 DNA.
4	2465.5	60.4	3924	24	Gene #2299 used to
5	2442	59.9	4279	22	Dog P-glycoprotein
6	2442	59.9	4279	22	Dog P-glycoprotein
7	2442	59.9	4279	22	Dog P-glycoprotein
8	2441	59.8	4186	22	Cynomolgous monke
9	2441	59.8	4195	22	Cynomolgous monke
10	2441	59.8	4279	22	Dog P-glycoprotein
11	2441	59.8	4369	21	Rat multidrug resi
12	2441	59.8	4425	21	Rat multidrug resi
13	2438	59.8	4317	22	Dog P-glycoprotein
14	2436	59.7	3840	24	Human polynucleoti
15	2436	59.7	3860	21	Human wild-type mu
16	2436	59.7	3860	21	Human G185V mutant
17	2436	59.7	3860	24	Human BCRP DNA rel
18	2436	59.7	3860	24	Human BCRP DNA rel
19	2436	59.7	3988	21	Human MDR-1 DNA.
20	2436	59.7	4264	19	Mutated human P-gl
21	2436	59.7	4264	19	Mutated human P-gl
22	2436	59.7	4349	22	Human intestine ce
23	2436	59.7	4646	21	Human ATP binding
24	2436	59.7	4646	24	Human mdr1 gene.
25	2436	59.7	4646	24	Kidney cancer rela
26	2436	59.7	4646	24	Kidney cancer rela
27	2436	59.7	4669	8	Sequence of human
28	2436	59.7	6505	17	Hybrid vector pSF-
29	2436	59.7	8630	21	Retroviral M4 mdr-
30	2436	59.7	8630	21	Retroviral vector
31	2433	59.6	4378	11	Multidrug Resistan
32	2433	59.6	4669	14	Sequence of human
33	2431.5	59.6	3912	24	Rat sequence of diffe
34	2425	59.5	4646	15	Human multidrug re
35	2425	59.5	4788	21	Murine multidrug r
36	2425	59.5	4788	24	Mouse BCRP DNA rel
37	2413.5	59.2	4233	21	Rat mdr1b2 multidr
38	2413.5	59.2	4233	22	Rat mdr1b2 multidr
39	2412	59.1	4669	19	Human P glycoprote
40	2412	59.1	4669	24	CDNA encoding huma
41	2404.5	58.9	4189	21	Murine multidrug r
42	2404.5	58.9	4189	24	Murine BCRP DNA rel
43	2404.5	58.9	4313	14	Mouse multidrug re
44	2372	58.2	4254	24	Rat sequence diffe
45	2280	55.9	2726	15	Multidrug-resistan

ALIGNMENTS

RESULT 1

ABK83223

ID ABK83223 standard; cDNA; 3699 BP.

XX

AC ABK83223;

XX

DT 27-AUG-2002 (first entry)

XX

DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CBI, cDNA.

XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;

KW neurological disorder; muscle disorder; immunological disorder; cancer;

KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;

KW cell proliferative disorder; cervical cancer; breast cancer;

KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;

KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;

KW Grave's disease; gastrointestinal disorder; Crohn's disease;

KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;



KW bacterial infection; fungal infection; parasitic infection;  
 KW protozoal infection; helminthic infection; cardiovascular disorder;  
 KW atherosclerosis; hepatic disease.

OS Homo sapiens.

PN WO200240541-A2.

PD 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

PR 03-NOV-2000; 2000US-245904P.

PR 17-NOV-2000; 2000US-247673P.

PR 20-NOV-2000; 2000US-249661P.

PR 01-DEC-2000; 2000US-252322P.

XX (INCY-) INCYTE GENOMICS INC.

PI Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

PI Walla NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

PI Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

PI Thangavelu K, Thornton M, Lu DM, Tribouley CM, Warren BA;

PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI: 2002-463570/49.

DR P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for  
 PT diagnosing, preventing, and treating disorders associated with an  
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular  
 PT or renal disorders

PS Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
 CC a recombinant polynucleotide comprising a promoter sequence operably  
 CC linked to the TRICH polynucleotide, a cell transformed with the  
 CC recombinant polynucleotide, a transgenic organism comprising the  
 CC recombinant polynucleotide, an isolated antibody that binds specifically  
 CC to TRICH, and screening for compounds which bind to TRICH, modulate  
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.

CC The polypeptides are useful for diagnosing, treating, and  
 CC preventing transport, neurological, muscle, immunological disorders  
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or  
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
 CC catatonla), endocrine disorders (e.g. diabetes, Grave's disease),  
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,  
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
 CC other diseases and disorders detailed in the specification. They can also  
 CC be used in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of transporters and  
 CC ion channels. TRICH or its fragments may also be used in screening for  
 CC compounds that specifically bind to and modulate the activity of TRICH.  
 CC The polynucleotides can be used to create knock-in humanised animals or  
 CC transgenic animals to model human disease. The present sequence  
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3699

Score: 3919.00 Matches: 785

Percent Similarity: 96.80% Conservative: 1

Best Local Similarity: 96.67% Mismatches: 0  
 Query Match: 96.08% Indels: 26  
 DB: 24 Gaps: 1

US-09-873-409-2 (1-812) x ABK83223 (1-3699)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
 Db 1339 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGCATATTCGAGACCATTATGGA 1398  
 QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIleTyrGly 40  
 Db 1399 GTGGTTAGTCAAGAGCCCTGTTTTTGTTCGGACCCACCATCAGTAAACAATATCAAGTATGGA 1458  
 QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
 Db 1459 CGAGATGATGTGACTGATGAGAGATGGAGAGAGCAGCAGGAGCAATCGGTATGAT 1518  
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
 Db 1519 TTTATCATGGAGTTTCTTAATAAATTTAATACATTTAGTGGGAAAAAGGAGCTCAATG 1578  
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProIleIle 100  
 Db 1579 AGTGAGGGCAGAAACAGAGGATCGCAATTCGTGCTTGTAGTTTCGAACCCCAAGATT 1638  
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
 Db 1639 CTGATTTTAGATGAGGCTACGTCTGCCCTGGANTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698  
 QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArgLeuSerThr 140  
 Db 1699 GCACCTGGAGAGGCGAGCAAGGTCGACTACAATCGTGTAGCACCCGACTTTCTACT 1758  
 QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
 Db 1759 ATTCGAAGTGCAGATTGTTGTCACCTAAAGGATGGAATGCTCGCGAGAAAGAGCG 1818  
 QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
 Db 1819 CATGCTGAACTAATGGCAAAAGAGTCTATATATTCCTTGTGATGTCACAGGATATT 1878  
 QY 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
 Db 1879 AAAAAGCTGATGAACAGATGAGTCAATGACATATCTACTGAAAGAAAGAACCAACTCA 1938  
 QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220  
 Db 1939 CTTCCTCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1998  
 QY 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240  
 Db 1999 ACCCAATCTAAAGAGATAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 2058  
 QY 241 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260  
 Db 2059 AAGCCTGAATGGCCCTTTGTGTTCTGGGACATTTGGCTCTCTTCTTAATGGAATGTT 2118  
 QY 261 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280  
 Db 2119 CATCCAGATATTTTCCATCATCTTTTCGAAAAAATTAACCATGTTTGGAAATAATGATAA 2178  
 QY 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleIleGlyValIle 300  
 Db 2179 ACCACATTAAGCATGATGCAGAAATTTATTCCAYGATATTTCGTCAATTTGGGTGTTATT 2238  
 QY 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320  
 Db 2239 TGCTTTGTTCAGTTATTTCATG----- 2259  
 QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340  
 Db 2260 -----CAGGATATTGCCTGGTTTGTAT 2280



	Qy	701	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAla	720
	Dd	3361	AACACAACAAGTTGGACTGAAGAGGACACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT	3420
	Qy	721	IleAlaARGAlaLeuLeuGlnLysProLysIleLeuLeuLeuASPgluAlaThr-SerAla	740
	Dd	3421	ATTGCCAAGGGCTCTTCTCAAAAACCCAAATTTATTGTGGATGAGCCACTTCAGCC	3480
	Qy	741	LeuASPAsnASerGlnLysValValGlnHisAlaLeuASPlysAlaArgThrGlyArg	760
	Dd	3481	CTCGATAATCACACAGTCAGAAGCTGGTTCAGCATGCCTTGATTAAGCCAGGACGGGAAGG	3540
	Qy	761	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaASPLeuValVal	780
	Dd	3541	ACATGGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCCAGATTGATGTGTT	3600
	Qy	781	LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgasp	800
	Dd	3601	CTGCACAATGGAAGATAAAGGAACAAGAACTCATCAAGAGCTCTCAGAAATCGAGAC	3660
	Qy	801	IleTyrPheLysLeuValAsnAlaGlnSerValGln	812
	Dd	3661	ATATATTAAAGTTAGTAGAATGCACAGTCAGTGCAG	3696
	RESULT	2		
	ID	AAZ94742	standard; cDNA; 3924 BP.	
	AC	AAZ94742;		
	XX			
	DT	01-AUG-2000	(first entry)	
	XX			
	DE	Human ATP binding cassette ABCB4 (MDR3) CDNA.		
	XX			
	KW	ABCB4; ATP binding cassette; human; cholesterol; lipid disorder; atherosclerosis; lipid disorder; dyslipidemia; psoriasis;		
	KW	lupus erythematosus; diagnosis; gene therapy; MDR3;		
	KW	multidrug resistance; chromosome 7q21; ss.		
	OS	Homo sapiens.		
	XX			
	FN	WO200018912-A2.		
	XX			
	PD	06-APR-2000.		
	XX			
	PF	21-SEP-1999; 99WO-EP06991.		
	XX			
	PR	25-SEP-1998; 98US-0101706.		
	XX			
	PA	(FARB ) BAYER AG.		
	XX			
	F1	Schmitz G, Klucken J;		
	XX			
	DR	WPI; 2000-293151/25.		
	XX			
	PT	Adenosine triphosphate binding proteins useful for identifying agents		
	XX	for treating atherosclerosis and other inflammatory disorders -		
	PS	Claim 9; Page 117-119; 154pp; English.		
	XX			
	CC	The present sequence is that of human ATP binding cassette		
	CC	subfamily B protein ABCB4 cDNA. The cDNA was identified using a		
	CC	differential display method in which monocytes from peripheral		
	CC	blood were subjected to macrophage differentiation and cholesterol		
	CC	loading with acetylated low density lipoproteins and subsequent		
	CC	deloading with high density lipoprotein (HDL3) to identify		
	CC	cholesterol sensitive genes. The gene maps to chromosome 7q21		
	CC	and is also termed MDR3 (multidrug resistance). The invention		
	CC	provides cholesterol-sensitive ABC genes (see AA294734-63). These		
	CC	genes, and polypeptides encoded by them, can be used for diagnostic		
	CC	and therapeutic applications, and for biochemical or cell-based		
	CC	assays to screen for pharmacologically active modulator compounds		

CC other inflammatory diseases such as psoriasis and lupus  
CC erythematosus.

SQ Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

**Alignment Scores:**

Pred. No.:	7.76e-229	Length:	3524
Score:	2465.50	Matches:	478
Percent Similarity:	76.4%	Conservative:	149
Best Local Similarity:	58.2%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	21	Gaps:	3

US-09-873-409-2 (1-812) x AAZ94742 (1-3924)

**Qy**      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
         ::::| | | | | | | | | | | | | | | | | | | | | |  
**D6**    1395 ATTGATGGCAGGATATTAGGAACCTTAATGTAAACTATCTGAGGGAAATCATTTGGTG 1454

Qy 22 valserGinGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1455 GTGAGTCAGAGACCGGTGCTTTTCCACCACAATTGCTGAAAAATATTTGTTATGGCGT 1514

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspSph 61  
                  :::|||||                  :::|||||                  :::|||||  
Db 1515 GGAAATGTAACCATGGATGAGATGAAGAAGCTGTCAAAGAGGCCAACGCCCTATGACTTT 1574

**Qy** 62 IleMetGluPheProAsnLysPheasnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
|||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||||::: |||||:::  
**Db** 1575 ATCATGAATATTACCA CAGAAATTTTGACACCCCTGGTGTGGAGAGAGAGGGGCCAGCTAGT 1634

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAargAlaLeuValArgasnProLysIleLeu 101  
 Db 1635 GGTGGGCGAAGCAGAGAGGATGCCATTGCAGTGCCTGGTTCGCACCCCAAGATCCTT 1694

**Qy** 102 IleLeuAspGluAlaThrSerAlaIeuApsSerGluSerLySerAlaValcInAlaAla 121  
:::|||||  
**Db** 1695 CTCTGGATGAGGCCACGTTCAGCATTTGGACACAGAAGTGAAGCTCAGGTACAGGCAGCT 1754

**DQ**

122 LeuGluLysAlaSerLysGlyArgThrIleValAlaHisArgLeuSerThrIle 141  
|||:::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
**DB**

1755 CTGGATAAGGCCACAGAAGCGCCGACCACCATTTGTATAGCACACCGACTGTCACGGTC 1814

**Qy** 142 ArgSerAlaAspLeuIleValThrLeuLysGlyMetLeuAlaGluLysGlyAlaHis 161  
|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||  
**Db** 1815 CGAAATGCAGATGTCATCGCTGGTGGTTTGAGATGGAGTAATTGTGGAGCAGGAAGCCAC 1874

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuVal----- 175  
:::||||||| | | | | | | | | | |  
Db 1875 AGCGAACTGATGAAGAGGAGGGGTACTTCAACTTGTCACATGCAGACATCAGA 1934

Qy 176 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192  
 Db 1935 AGCCAGATCCAGTCAGAAGAAATTTGAACCTAAATGATGAAAAGGCTGCCACTAGAAATGGCC 1994

Qy 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211  
|||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::  
Db 1995 CCAAATGGCTGGAATCTCGCCTATTAGGCATTCTACTCAGAAAAACCTTAAATAATCA 2054

Qy 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysLysGluIleSerLeuProGlu 230  
Db 2055 CAAATGTGTCAAGAGACCTTGATGTGGAAACCGATGACCTTGAGCAAAATCTGCCACCA 2114

Qy	231	ValSerLeuLeuLysIleLeuLysLysProGluTrpProPheValValLeuGly	250
		:::	
			:::
Db	2115	GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAAATGSCCCTACTTTCTCGTGGGA	2174

Qy	251	ThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIleIlePheAlaLys	270
Db	2175	ACAGTATGTGCCATGTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTTCTCAGAG	2234

QY 271 ILeileThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290  
||||| :||| ||| :: |||::

Db	2235	ATCATAGCCATTTTGGACCGCCATCATGCATGAAGCAGCAGACAGTGCACATATTC	2239
Qy	291	SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu	310
Db	2295	TCITTGATTTTCTTATTTCTGGGAATATTTCTCTTTTACTTCTTCTCTCCAGGTTTC	2354
Qy	311	PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla	330
Db	2355	ACGTTTGGGAAAGCTGGCGAGATCTCCACGAAGACATGCGGTCAATGGCTTTTAAAGCA	2414
Qy	331	MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr	350
Db	2415	ATGCTAAGACAGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGTGCACATTCT	2474
Qy	351	ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal	370
Db	2475	ACAAAGACTTGGCCACAGATGCTCCCAAGTCCCAAGAGGCCACAGAAACCGGTTGGCTTTA	2534
Qy	371	LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr	390
Db	2535	ATTGCACAGAATATAGTCAACCTTGGAACTGGTATTATCATCATTTATCTACGGTTGG	2594
Qy	391	GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu	410
Db	2595	CAGTTAAACCTATTGTATTAGCAGTTGTTCCAAATATTGCTGTGTCAGGAATTTGTGAA	2654
Qy	411	ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLys	430
Db	2655	ATGAAATTGTTGGCTGGAAATGCCAAAGAGATAAAAAGAACTGGGAAGCTGGTGAAG	2714
Qy	431	IleAlaThrGluAlaLeuGluAsnIleAlaThrIleValSerLeuThrArgGluLysAla	450
Db	2715	ATTGCAACAGAGGCAATAGAAATATTAGGACAGTTGTCTTTTGAACCCAGCAAGAAAA	2774
Qy	451	PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla	470
Db	2775	TTTGAATCAATGTATGTGCAAAATTTGTATGGACCTTACAGGAATTTCTGTCAGAAAGCA	2834
Qy	471	GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla	490
Db	2835	CACATCTATGGAATTACTTTTAGTATCTCCACAGCAATTTATGTATTTTCTATGCCGT	2894
Qy	491	GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe	510
Db	2895	TGTTTTCGATTTGGTCATATCTCATTTGGAATGGACATATGCGCTTCAGAGATGTTATT	2954
Qy	511	IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla	530
Db	2955	CTGGTGTTTTCTGCAATTTGATTTGGTCAGTGGCTCTAGACATATCCAGTTCATTGCT	3014
Qy	531	ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys	550
Db	3015	CCAGACTATGCTAAAGCTAAGCTGTCTGCAGGCCACTATTATCATGCTGTTTGAAGAAGCA	3074
Qy	551	ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu	570
Db	3075	CTCTGATTTGACAGCTACAGTGAAGGGGCTGAAGCCCTGATAAATTTGAAGGAAATATA	3134
Qy	571	GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly	590
Db	3135	ACATTTAATGAAGTCGTGTTCAACTATCCACCCGAGCAAAACGTGCCACGTCTTCAGGGG	3194
Qy	591	LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly	610
Db	3195	CTGAGCCTGGAGGTGAAGAAAGCCAGACACTAGGCCCTGGTGGGCGAGCATGGCTGTGGG	3254
Qy	611	LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu	630
Db	3255	AAGACACGGTGTCTCAGCTCTCTGGAGCGGTCTTACGACCCCTTGGCGGGGACAGTGCTT	3314
Qy	631	PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle	650
Db	3315	CTCGATGGTCAAGACCAAGAAATCAATGTCCAGTGGCTCAGAGTCAATCCGGAATC	3374



QY 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211  
DB 1995 CCAAAATGGCTGGAAATCTCGCCTATTAGGCATTTCTACTCAGAAACAACTTAAATAATCA 2054  
QY 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230  
DB 2055 CAAATGTGTGAGAGAGCTTGAATGTGGAAACCGATGGACTTGAACCAATGTGCCACCA 2114  
QY 231 ValSerLeuLysLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250  
DB 2115 GTGTCCTTTCTGAAGTCTGAACTGAATAAAACAGAAATGCCCTACTTTGTGCTGGGA 2174  
QY 251 ThrLeuAlaSerValLeuAsnGlnThrValHisProValPheSerIleIlePheAlaLys 270  
DB 2175 ACAGTATGTGCCATTTGCCAATGGGGGCTTCAGCCGGCATTTTTCAGTCATATTTCTCAGAG 2234  
QY 271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290  
DB 2235 ATCATAGCATTTTGGACCGCGATGATGTCAGTGAACGAGAGAGTGCACATATTTCT 2294  
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310  
DB 2295 TCTTTGATTTCTTATTCTGGGAATTTTCTTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2354  
QY 311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330  
DB 2355 ACGTTTGGGAAGCTGGCGAGATCTCCACAGAGAGACTCGCGTCAATGGCTTTTAAAGCA 2414  
QY 331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350  
DB 2415 ATGCTTAAGCAGACAGTACGTGTTTGTGATGACCAATAAAACAGTACTGCTGCATTTCT 2474  
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370  
DB 2475 ACAAGACTTGGCCACAGATGCTGCCAAGTCCAAGGCCACAGGACCAAGTGGCTTTA 2534  
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390  
DB 2535 ATGTCACAGAAATATAGCTAACCTTGAACCTGGTATTTATCATATCATTTATCTACGGTTGG 2594  
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410  
DB 2595 CAGTTAACCTTATGCTATTAGCATGTTTCCAAITATTGCTGTGTCAGGAATGTTGAA 2654  
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430  
DB 2655 ATGAATTTGTGCTGGAATGCTCAAAAGAGATATAAAAGAACTGGAAGCTGCTGGAAG 2714  
QY 431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450  
DB 2715 ATTGCAACAGAGCAATAGAAAATATTAGACAGTTGTGCTTTCAGCCAGGAAAGAAA 2774  
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470  
DB 2775 TTGTAATCAATGATTTGAAAATTTGATGACCTTACAGCAATTTCTGTGAGAGGCA 2834  
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490  
DB 2835 CACATCTATGGAATTTTATGATCTCACAAGCATTTATGATTTTCTTATGCGGT 2894  
QY 491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510  
DB 2895 TGTGTTTCGATTTTGTGCATATCTCATTTGTAATGGACATATGGCTTCAGAGATGTTATT 2954  
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530  
DB 2955 CTGGTGTGTTTCTGCAATTTGATTTGTGTCAGTGGCTTAGGACATGCCAGTTCATTGCT 3014  
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 550  
DB 3015 CCAGCATGCTAAAGCTGCTGAGCCCACTTATTCATGCTGTTTGAAGACAA 3074  
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570

DB 3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGTGAAGCCCTGATAAAATTTTGAAGAAATATA 3134  
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590  
DB 3135 ACATTTAATGAAGTCTGTTCACATATCCACCCGAGCAACGTCAGTGTCTCAGGGG 3194  
QY 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610  
DB 3195 CTGAGCTGGAGTGAAGAAAGCCAGACACTAGCCCTGGTGGCAGCAGTGCCTGTGGG 3254  
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeu 630  
DB 3255 AAGACACGCTGGTCCAGCTCTGAGCGGTTCACGACCCCTTGGCGGGGACAGTGT 3314  
QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650  
DB 3315 CTCGATGTCAAGAACCAAGAACTCAATGTCCAGTGGCTCAGAGCTCACTCGAATC 3374  
QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670  
DB 3375 GTGCTCAGGAGCTTATCTATTGCTGACCTGACATTTGCCAGAAATTTGCTTATGGAGAC 3434  
QY 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHis 690  
DB 3435 AACACCCCGGTGTATCATCAGGATGAAATTTGTAGTGCAGCCCAAGCTGCCAACATACAT 3494  
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710  
DB 3495 CTTTTCATCGAGACCTTATCCCAACAATATGAACAGAGTGGGAGATAAGGGACTCAG 3554  
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730  
DB 3555 CTCTCAGGAGTCAAAAACAGAGGATGCTATTGCTGCCAGCCCTCATCAGACAACTCAA 3614  
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750  
DB 3615 ATCTCTCTGTTGGATGAAGCTACATCAGCTCTGGATCTGAAAGTGAAGAGTGTCCAA 3674  
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770  
DB 3675 GAAGCCCTGGACAAAGCCAGAGAGCCGACCTGATGTTGATGCTCACCCCTGCTCC 3734  
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790  
DB 3735 ACCATCCAGAAATGCAGACTTAATATGCTGTTTCAGAAATGGAGAGTCAAGGAGCATGGC 3794  
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810  
DB 3795 ACGCATCAGCAGCTCTGCGCAGAAAGGCATCTATTATTTTCAATGCTCAGTGTCCAGGCT 3854  
RESULT 4  
ABN95801  
ID ABN95801 standard; DNA; 3924 BP.  
XX  
AC ABN95801;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE Gene #2299 used to diagnose liver cancer.  
XX  
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;  
metastatic liver tumour; cytostatic; expression profile; disease state;  
disease progression; drug toxicity; drug efficacy; drug metabolism.  
XX  
OS Homo sapiens.  
XX  
PN WO200229103-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 02-OCT-2001; 2001WO-US30589.  
XX

PR 02-OCT-2000: 2000US-237054P.

XX  
PA (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG; PI

XX  
DR WPI; 2002-426119/45.

XX  
PT  
Diagnosing and detecting the progression of liver cancer,

PT hepatocellular carcinoma or metastatic liver tumor in a patient, PT involves detecting the level of expression of two or more genes

PT - liver tissue sample  
XX

PS Claim 1; SEQ ID NO 2299; 298pp; English.

XX

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in AEN93503-AEN97455 in a tissue sample. The method of the invention has hepatotropic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct](http://ftp.wipo.int/pub/published/pct) sequences.

Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:		
Pred. No.:	7,766-229	Length:
Score:	2465.50	Matches:
Percent Similarity:	76.4%	Conservative:
Best Local Similarity:	58.2%	Mismatches:
Query Match:	60.4%	Indels:
DB:	24	Gaps:
		3
		3924

US-09-873-409-2 (1-812) x ABN95801 (1-3924)

Qy	2	ValAspGluAenAspIleA	ArgAlaLeuAenValA	ArgHisTyrA	ArgAspHisIleGly	Val	21
		::::	::::	::::	::::	::::	
Db	1395	ATTGATGGG	CAGGATATTAG	GAACCTTAA	TGTAACCTATCT	GAGGGAATCAT	TTGGTG
		::::	::::	::::	::::	::::	
Qy	22	ValSerGlnGluProVal	LeuPheGlyThr	ThrIleSerAsn	AsnIleGlyTyrGly	Arg	41
		::::	::::	::::	::::	::::	
Db	1455	GTGAGTCAGG	AGCGGTCTG	TTTCCACCA	CAATGCTG	GAATAATTT	TGTTATGGCCGT
		::::	::::	::::	::::	::::	
Qy	42	AspAspValThrAsp	GluGluMetGluA	ArgAlaA	ArgGluAlaAsn	AlaTyrAsp	Phe
		::::	::::	::::	::::	::::	
Db	1515	GGAAATGAA	CCATGGATGAG	ATAAGAAAGCTG	CAAAAGAGGCCA	ACCGCTAT	GAGTTT
		::::	::::	::::	::::	::::	
Qy	62	IleMetGluPhePro	AnIysPheAenThr	LeuValGlyGlu	LysGlyAlaGln	MetSer	81
		::::	::::	::::	::::	::::	
Db	1575	ATCATGAAT	TACACAGAAAT	TTGACACCTG	GTGTGGAGAGAGGG	CGCCAGCT	GAGT
		::::	::::	::::	::::	::::	
Qy	82	GlyGlyGlnLysGln	ArgIleAlaIleA	ArgAlaLeuValA	ArgAsnProLysIle	Leu	101
		::::	::::	::::	::::	::::	
Db	1635	GTTGGG	CAGAAGCAGAGG	ATCGCATTTG	CACGTGCGCTG	TTTCGCAACCC	CAAGATCCTT
		::::	::::	::::	::::	::::	
Qy	102	IleLeuAspGluAla	ThrSerAlaLeuA	AspSerGluSerLys	SerAlaValGln	AlaAla	121
		::::	::::	::::	::::	::::	
Db	1695	CTGTGTGAT	CAGGCCACGT	CAGCATTTG	SGACACAGAAAGTGA	AGCTGAGCGT	ACAGCGAGCT
		::::	::::	::::	::::	::::	
Qy	122	LeuGluLysAlaSer	LysGlyArgThr	ThrIleValValAla	HisArgLeuSerThr	Ile	141
		::::	::::	::::	::::	::::	
Db	1755	CTGGAT	AAGGCCAGAG	AGCCCGGACCA	CCATTGTGTATG	ATGACACCG	ACTGTCTACGGTC
		::::	::::	::::	::::	::::	
Qy	142	ArgSerAlaAsp	LeuIleValThr	LeuLysAspGly	MetLeuAlaGlu	LysGlyAla	His
		::::	::::	::::	::::	::::	

```
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 2955 CTGGTGTTCGCAATGTATTGGTCAGTGGCTTAGACATCCAGTTCATTGCT 3014
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 550
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGGCCACTTATTCATGCTGTTGAAAGCAA 3074
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3075 CCTCTGATTGACAGCTACAGTGAAGAGGGCTGAAGCCCTGATTAATTTGAAGAAATATA 3134
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 571 GluPheArgGluValSerPheThrProCysArgProAspValPheIleLeuArgGly 590
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3135 ACATTTAATGAAGTCGTGTTCAACTATCCACCAGCAAGCAACGTCAGTCTTCAGGG 3194
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 591 LeuSerIleuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGCGACAGTGGCTGTGG 3254
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3255 AAGAGCAGGTGTCTCAGCTCTGGAGCGGTTCTACGCCCTTGGCGGGACAGTGT 3314
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3315 CTGATGGTCAAGCAAGCAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3375 GTCTCTCAGAGCCTATCTCTATTGATGAGTGAATTTGTGAGTGCAGCCAAAGCTGCAACATACAT 3434
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaIleHis 690
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3435 AACAGCGGGGTTGTATCAGAGTGAATTTGTGAGTGCAGCCAAAGCTGCAACATACAT 3494
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGln 710
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3495 CCTTTTCATCAGACGTTACCCCAATATGAACAGAGTGGAGATGAAGGGATCTAG 3554
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuLysProLys 730
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATTTGTTATTCGCCGAGCCCTCATCAGAACCTCAA 3614
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3615 ATCCTCTGTGTGATGAAGCTACATCAGCTCTGGATCTGAAGTGAAGAAAGGTGTGCCAA 3674
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3675 GAAGCCCTGGCAAAAGCCAGAGAGCCGACCTGTCATTGTGATTCCTACCGCCTGTCC 3734
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 771 AlaIleGluAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3735 ACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAGAAATGGGAGAGTCAAGGAGCATGGC 3794
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 791 ThrHisGlnLeuLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3795 ACGCATCAGCAGCTGTGGCACAGAAAGGCATCTATTTTTCATGTGTCAGTGTCCAGGCT 3854
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
RESULT 5
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
  drug bioavailability; transgenic animal; genetic model; ss.
```

```
XX
OS Canis familiaris.
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
PN WO200123540-A2.
PD 05-APR-2001.
PF 28-SEP-2000; 2000WO-US26767.
PR 28-SEP-1999; 99US-0156510.
PA (GENT-) GENTEST CORP.
XX
XX Stocker PU, Steimel-crespi DT, Crespi CL, Reif TC, Patten CU;
PI P-PSDB; AAE00303.
DR WPI; 2001-235373/24.
DR P-PSDB; AAE00303.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 3; Page 58-63; 111pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity. By
CC reducing or increasing PGP activity in a cell, PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.: 1,69e-226 Length: 4279
Score: 2442.00 Matches: 469
Percent Similarity: 76.24% Conservative: 160
Best Local Similarity: 56.85% Mismatches: 178
Query Match: 59.87% Indels: 18
Dbs: 22 Gaps: 4

US-09-873-409-2 (1-812) x AAD03488 (1-4279)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1379 ATTGATGGACAGACATTTAGGACCATAAATGTAAGCATCTTCGGGAAATTAATCTGGTGTG 1438
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1439 GTGAGTCAGGAGCCTGTGTGTTTCCACCACGATAGCTGTAAGAAACATTCGCTATGCGCGC 1498
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1499 GAAATGTCACCATGATGAGATTGAGAAAGCTGTTAAGAGAGCCATGCTATGATTTT 1558
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```



Db 1559 ATCTGAAACTACCTAATAAATTTGACACTCTGGTGGAGAGAGGGGCCAGCTGAGT 1618  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
Db 1619 GTGTGACAGAAACAGAGAAATGCCATTTGCTCGGGCCCTGGTTGCGAAACCCCAAGATTCTT 1678  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLa 121  
Db 1679 CTGCTGGATGAGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCCAGGTGGCC 1738  
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 1739 CTGGATAGGCCAGAAAGCCGCGACTACCATTTGTATGATAGCTCATCGTTTGTCTACAGTT 1798  
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
Db 1799 CGTAATGCCGATGTCATTCTGCTGGTTTGTATGATGGATGTCATTTGTGGAGAAAGAAATCAT 1858  
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyfTyfSerLeuValMetSerGlnAspIleLys 181  
Db 1859 GATGAACCTCATGAAAGAGAGGCAATTTACTTCAAACTTGTCACAATGCGAG--ACAAAGA 1915  
Qy 182 LysAlaAspGluGlnMetGluSerMetThrTyfSerThrGluArgLysThrAsnSerLeu 201  
Db 1916 GGAAATGAATGAGTTAGTAANAATGCCACTGGTGAATCCANAAGTGAAAGTGAATGCTCTTG 1975  
Qy 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214  
Db 1976 GAAATG---TCTCCAAAGATTCAGGTCAGTTTATTAATAAAGAAAGATCAACTCGCAGG 2032  
Qy 215 -----AspLysAlaGluLysSerThrGlnSerLysGluIle 226  
Db 2033 ACTATACATGCAACCAAGGCCAAGACAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2092  
Qy 227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246  
Db 2093 AATGTACCTCCAGTTTCCCTTCCTCGAGGATTCGAGCTGAACCTGAACCTGAATGGCCTTAT 2152  
Qy 247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266  
Db 2153 TTTTGGTTGGTATATTTTGTCTATTATAAAGCGAGGCTCGAACACACATTTTCAATA 2212  
Qy 267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285  
Db 2213 ATATTTTCAAGGATTATAGGATCTTTACCCGAGATGAGGATCCTGGAACAAACAGACAG 2272  
Qy 286 AspAlaGluIleTyfSerMetIlePheValIleLeuGlyValIleCysPheValSerTyf 305  
Db 2273 AATAGTAACATGTTTCTGTATTGTTTCTAGTCTTGGAAATATTCTTTTATTACATTT 2332  
Qy 306 PheMetGlnGlyLeuPheTyfGlyArgAlaGlyGluIleLeuMetArgLeuArgHis 325  
Db 2333 TTTCTCCAGGGCTTCATTTTGGCAAGCTGGGGAGATCCTCACTAAGCGGCTTGATAC 2392  
Qy 326 LeuAlaPheLysAlaMetLeuTyfGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345  
Db 2393 ATGGTTTTCAGATCCATGCTGACACAGGATGTCAGCTGGTTTGTATGATCACCTTAAACACC 2452  
Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365  
Db 2453 ACTGGAGCATTCACAAACAGGCTTGCCATGATGCGGCTCAAGTTAAAGGGGCTATAGT 2512  
Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385  
Db 2513 TCCAGCTTGGCTGTCATTATCCCAAGATATAGCAAAATCTTTGGGACAGGCAATTATATATCC 2572  
Qy 386 PheIleTyfGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405  
Db 2573 TTAATCTATGGTTGGCAATTAACACTTTTACTTCTAGCAATTTGTACCCATCATTTGCAATA 2632  
Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425  
Db 425

Db 2633 GCAGGAGTTGTTGAAATGAAATGTTGTCTGACACAGCACTGAAGAAGATAAGAAAGAGCTA 2692  
Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445  
Db 2693 GAAGGAGCTGGAGAGATTGCTACAGAACCCATCGAAAACTTCCGAACACTGTTGTTCTTTG 2752  
Qy 446 ThrArgGluLysAlaPheGluGlnMetTyfGluGluMetLeuGlnThrGlnHisArgAsn 465  
Db 2753 ACTCGGAGCAGAAAGTTTGAATATACATGTATGCACAGAGTTTGCACAGTACCATACAGAAAC 2812  
Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyfAlaPheSerHisAlaPheIleTyf 485  
Db 2813 TCTTTGAGAAAGACACATCTTCGGGTCTCATTTTATCACCAGCAATGATGAT 2872  
Qy 486 PheAlaTyfAlaAlaGlyPheArgPheGlyAlaTyfLeuIleGlnAlaGlyArgMetThr 505  
Db 2873 TTTTCTATGCTGGCTGTTTCCGGTTTGGTCTACTTGGTGGCAATGAGTTTCATGAAC 2932  
Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyfGlyAlaMetAlaIleGlyLys 525  
Db 2933 TTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGCTTTGGTGCATGCGCAGTGGGCGAG 2992  
Qy 526 ThrLeuValLeuAlaProGluTyfSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545  
Db 2993 GTCAGTTTCATTTGCTCTCTGACTATGCCAAAGCCAAAGTATCAGCAGGCCACGTCATCATG 3052  
Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565  
Db 3053 ATCATTTGAAAGAGCCCTCTGATTGACAGCTACAGCCCTCAGGCTCAAGCCAAATACG 3112  
Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyfProCysArgProAspVal 585  
Db 3113 TTGGAAGCAATGTGACATTTAATGAGTCTGTTTCAACTATCCCACTCGCACACACATC 3172  
Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605  
Db 3173 CCCGTCTCCAGGGCTGAGCCTCAGGTGAAGAGGCGCACACGCTGCCCTCTAGTAGT 3232  
Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyfAspProVal 625  
Db 3233 AGCATGGCTGTGGAGAGACACATGTTCTTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292  
Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645  
Db 3293 GCTGCTTCAGTGTAAATTCATGGCAAGAGATAAAGACACCTGAATGTCCAGTGGCTCGGA 3352  
Qy 646 SerGlnIleAlaIleValProGlnclProValLeuPheAsnCysSerIleAlaGluAsn 665  
Db 3353 GCACACCTGGGGCATCGTGTCTCAGAGCCCATCTCTGTTGACTGCGACATTCGCCAGAAC 3412  
Qy 666 IleAlaTyfGlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsn 685  
Db 3413 ATTGCTATGAGACACACAGCCGGTCTGATCATCAATGAAGAGATTATGCAGGCGACCAAG 3472  
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyfAsnThrGlnValGly 705  
Db 3473 GAGGCCAACATACACACTTCATCAGACACTCTCCCTGAGAAATACAAACACAGAGTAGGA 3532  
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725  
Db 3533 GACAAAGAACCCAGCTCTCTGGTGGCAGAAACAGCCATTCGCCATAGCTCGCGCTCTT 3592  
Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745  
Db 3593 GTTAGACAGCTCATATTTTGTCTTTGGATGAAGTACATCAGCTCTCGATACAGAAAGT 3652  
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765  
Db 3653 GAAAGGTTGTCAGAAAGCCCTGGCAAAAGCCAGAGAAAGCCGACCTGTCATTTGATGATC 3712  
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785  
Db 3713 GCCCACCGCTTGTCCACCATCAGAAATGCAGATTTAATAGTGTGTTTTCAGAAATGCAGAA 3772

QY 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805  
 Db 3773 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCAGAAAGGCATCTATTTTCCATG 3832

QY 806 ValAsnAlaGlnSer 810  
 Db 3833 GTCAGTGTCCAGGCT 3847

RESULT 6  
 AAD03504  
 ID AAD03504 standard; cDNA; 4279 BP.  
 XX  
 AC AAD03504;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.  
 XX  
 KW Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;  
 KW drug bioavailability; transgenic animal; genetic model; ss.  
 XX  
 OS Canis familiaris.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 17..3862  
 FT /\*tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype A) protein"  
 FT replace (607, C)  
 FT /\*tag= b

WO200123540-A2.  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26767.  
 XX  
 PR 28-SEP-1999; 99US-0156510.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
 XX  
 DR WPI; 2001-235373/24.  
 DR P-PSDB; AAE00308.  
 XX  
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 PS Claim 9; Page 85-90; 11pp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 XX  
 SQ Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.69e-226 Length: 4279  
 Score: 2442.00 Matches: 469  
 Percent Similarity: 76.24% Conservative: 160  
 Best Local Similarity: 56.85% Mismatches: 178  
 Query Match: 59.87% Indels: 18  
 Ds: 22 Gaps: 4

US-09-873-409-2 (1-812) x AAD03504 (1-4279)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1379 ATTGATGACAGGACATAGGACCAATAATGTAGGCATCTTCGGGAAATACTCGTGTG 1438

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1439 GTGACTCAGGAGCCTGTGTGTTGCCACACGATAGCTGAAACATTCGTATGCGCGC 1498

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 Db 1499 GAAATGTCCACATGATGAGATTGAGAAAGCTGTTAAGGAAGCCATGCTATGATTTT 1558

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1559 ATCATGAAACTACCTAATAAATTTGACACTCTGTTGGAGAGAGGCGCCAGCTGAGT 1618

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1619 GGTGACAGAAACAGAGAAATGCCATTGCTGGGCCCTGGTTGCGAACCCCAAGATTCTT 1678

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 121  
 Db 1679 CTGCTGGATGAGGCAACGTCAGCTCTGCACACTGAAAGTGAAGCAGTGGTTCAGTGCC 1738

QY 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaHisArgLeuSerThrIle 141  
 Db 1739 CTGGATAAGGCCAGAAAGCGCGGACTACCATTTGATGCTCATCGTTGTCTACAGTT 1798

QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
 Db 1799 CGTAATGCCGATGTCATTGCTGGTTTTCATGATGGAGTCATGTGGAGAAAGGAATCAT 1858

QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181  
 Db 1859 GATGAACATCATGAAAGAGAGGCGCATTTACTTCAAACTGTGCACAAATGCAG---ACAAGA 1915

QY 182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201  
 Db 1916 GGAATGAAATTGAGTTAGAAATGCCACTGGTGAATCCAAAGTGAAAGTGAATGCGCTTG 1975

QY 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214  
 Db 1976 GAAATG---TCTCCAAAGATTTCAGGTCAGTTTAATAAAGAGAGATCACTCGCAGG 2032

QY 215 -----AspLysAlaGluGluSerThrGlnSerLysGluIle 226  
 Db 2033 AGTATACATGCACCAAGCGGCAAGAGAAAGCTGTGTACAAAAGAGGACTTGAATGAG 2092

QY 227 SerLeuProGluValSerLeuLysIleLysLeuAsnLysProGluTrpProPhe 246  
 Db 2093 AATGTACCTCCAGGTTTCTTCTGGAGGATTCGAAAGCTGAACTCAACTGAATGCGCTTAT 2152

QY 247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266  
 Db 2153 TTTGGTGGTGTATATTTTGTGTATATAACCGAGGCGCTGCAACCAAGCATTTCAATA 2212

QY 267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285  
 Db 2213 ATATTTTCAAGGATATAGGAGTCTTTTACCAGAGATGAGGATCTCTGAAACAAACACGAG 2272

QY 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305  
 Db 2273 AATAGTAACATGTTTCTGTGTTGTTCTAGTCTCTGGAATATTATTCTTTTATACATT 2332



Qy 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluileLeuThrMetArgLeuArgHis 325  
 Db 2333 TTCTCCAGGGCTTCCATTTGGCAAGCTGGGAGATCTCTCACTAAGCGGCTCGATAC 2392  
 Qy 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTtpPheAspGluLysGluAsnSer 345  
 Db 2393 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCACTGGTTGTAGCCCTAAACACACC 2452  
 Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365  
 Db 2453 ACTGAGCATTTGACACACAGGCTTGCCATGATCGCGCTCAAGTTAAAGGGGCTATAGT 2512  
 Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385  
 Db 2513 TCCAGGCTTGCTGCTATTACCCAGCAATATAGCAATCTTGGGACAGGCATTTATATATCC 2572  
 Qy 386 PheIleTyrGlyTtpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405  
 Db 2573 TTAATCTATGTTGGCAATTAAACATTTTACTCTTAGCAATTTGTACCCATCAATGCAATA 2632  
 Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425  
 Db 2633 CGAGAGTTGTTGAAATGAAATGTTGCTGCAAGCACTCAAGATTAAGAAAGCTA 2692  
 Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445  
 Db 2693 GAAGAGCTGGGAAGATTGCTACAGAAGCCATCGAAACCTTCGAACTGTTGTTCTTTG 2752  
 Qy 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465  
 Db 2753 ACTCGGAGCAGCAAGTTTGAATACATGATGTATGCACAGAGTTTGCAGATACCATACAGAAC 2812  
 Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485  
 Db 2813 TCTTTGAGGAAGACACATCTTCGGGGTCTCATTTCTATCACCCAGCAATGATGAT 2872  
 Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505  
 Db 2873 TTTTCTATGCTGGCTGTTTCGGGTTTGGTCTACTTGGTGCAAAATCAGTTTCATGAAC 2932  
 Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525  
 Db 2933 TTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGCTTTGGTGCCATGTCAGTGGGCGAC 2992  
 Qy 526 ThrLeuAlaLeuAlaProGluTyrSerLysAlaLysSerGlyValAlaAlaHisLeuPheAla 545  
 Db 2993 GTCAGTTCAATTTGCTCTGACTATGCAAGCCAAAGTATCAGCAGCCCAAGTCATCATG 3052  
 Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565  
 Db 3053 ATCATTTGAAAAGCCCTCTGATTGACAGTACAGCCCTCAGCGCTCAAGCCAAATACG 3112  
 Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585  
 Db 3113 TTGGAAGGAATGTCACATTTAATGAGTCTGTTCAACTATATCCCACTCGACAGACATC 3172  
 Qy 586 PheIleLeuArgGlySerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605  
 Db 3173 CCGGTCTCAGGGGCTGAGCCTCGAGGTTGAAGAGGGCCAGACGCTGGCCCTCGTAGGT 3232  
 Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625  
 Db 3233 AGCAGTGGCTGTGGGAAGACACAGTTGTTACGTTCTTAGAGCGCTTCTATGACCCCTTG 3292  
 Qy 626 GlnGlyGlnValLeuPheAspGlyValAlaPheAlaLysGluLeuAsnValGlnTyrLeuArg 645  
 Db 3293 GCTGGTTCAAGTCTAATGATGGCAAGAGATTAAGCACCTGAATGTCAGTGGCTCCGA 3352  
 Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665  
 Db 3353 GCACACCTGGGCTATGTTCTCAGGAGCCCATCTGTTTGAAGTGCAGCATTCGCCGAGAAC 3412

Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685  
 Db 3413 ATTCCCTATGAGACCAACAGCCGGTCTGATTCATGAAGAGATTATCGAGCGCAGCCAAAG 3472  
 Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705  
 Db 3473 GAGGCCAATACATACACCACTTCATCGAGACACTCCCTGAGNAATACCAACACAGAGTAGGA 3532  
 Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725  
 Db 3533 GACAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTTGCCATAGCTCGCGCTCTT 3592  
 Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745  
 Db 3593 GTTACAGACCTCATATTTTCTTTGGATGAAGCTACATCAGCTCTCGATACAGAAAGT 3652  
 Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765  
 Db 3653 GAAAAGTTGTTCCAAAGCCCTGGACAAAGCCAGAAAGCCGACCTGTCATTTGTGATC 3712  
 Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785  
 Db 3713 GCCCACCCTTGTCCACCATCCAGATTCAGATTTAATAGTGGTGTTCAGAAATGGCAA 3772  
 Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805  
 Db 3773 GTCAGGAGCATGGCACATCAACAGCTGCTGGCCCAAGAGGCATCTATTTTTCATG 3832  
 Qy 806 ValAsnAlaGlnSer 810  
 Db 3833 GTCAGTGTCCAGGCT 3847  
 RESULT 7  
 ID AAD03505 standard; cDNA; 4279 BP.  
 AC AAD03505;  
 DT 13-JUN-2001 (first entry)  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.  
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.  
 OS Canis familiaris.  
 FH Key Location/Qualifiers  
 CDS 17..3862  
 FT /\*tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype B) protein"  
 FT replace (91, T)  
 FT /\*tag= b  
 FT replace (607, C)  
 FT /\*tag= c  
 PN WO200123540-A2.  
 XX  
 PD 05-APR-2001.  
 PF 28-SEP-2000; 2000WO-US26767.  
 PR 28-SEP-1999; 99US-0156510.  
 XX (GENT-) GENTEST CORP.  
 XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
 DR WPI; 2001-235373/24.  
 DR P-PSDB; AAE00309.  
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT

PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -

Claim 9; Page 93-99; 11pp; English.

CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.

XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,69e-226	Length:	4279
Score:	2442.00	Matches:	469
Percent Similarity:	76.24%	Conservative:	160
Best Local Similarity:	56.85%	Mismatches:	178
Query Match:	59.87%	Indels:	18
DB:	22	Gaps:	4

US-09-873-409-2 (1-812) x AAD03505 (1-4279)

QY	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisThrArgAspHisIleGlyVal	21
DB	1379	ATTGATGGACAGACATTAAGACCAATAAATGTAAAGCATCTTCGGAAATTAATCGTGTG	1438
QY	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
DB	1439	GTGAGTCAGAGGCTGTGTGTTGGCCACCATGATAGCTGAAACATTCGCTATGCGCCG	1498
QY	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaThrAspPhe	61
DB	1499	GAATAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAGCCCAATGCCATGATTTT	1558
QY	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
DB	1559	ATCATGAAACTACTTAATAAATTTGACACTCTGTGTTGGAGAGAGAGGGGCCAGCTGAGT	1618
QY	82	GlyGlyGlnLysGlnArgIleAlaIleAlaAatGAlaLeuValArgAsnProLysIleLeu	101
DB	1619	GGTGGACAGAAACAGAGAAATCGCCATTCGTGGGCCCTGTGTCGCAACCCCAAGATTC	1678
QY	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
DB	1679	CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAACGACGTGTTTCAGGTGGC	1738
QY	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgIleuSerThrIle	141
DB	1739	CTGGATTAAGCCAGAAAGCCGAGCTACCATTTGTGATAGCTCATCGTTTGTCTACAGTT	1798
QY	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
DB	1799	CGTAATGCCATGATCATTCGTGTTTGTATGATGGAGTCAATGTGGAGAAAGAAATCAT	1858
QY	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys	181
DB	1859	GATGAACATCATGAAGAGAGAGGCGCAITTACTTCAAACTTGTCAATATGCAG---ACAAGA	1915
QY	182	LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu	201
DB		::: ::::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	

DB	1916	GGAAATGAAATTTAGTTAGTAAATGCCACTGGTGAATCCAAAGATGAAAGTGAATGCTG	1975
QY	202	ProLeuHisSerValLysSerIleLysSerAspPheIle-----	214
DB	1976	GAAATG---TCTCCAAAGATTCAGGGTCCAGTTTAATAAATAAAGATCAACTCGCAGG	2032
QY	215	-----AspLysAlaGluGluSerThrGlnSerLysGluIle	226
DB	2033	AGTATACATGCACCAAGGGCCAGCAAGAAAGCTTGGTACAAAGAGAGACTTGAATGAG	2092
QY	227	SerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe	246
DB	2093	AATGACCTCCAGTTCTCTCGAGGATCTGAAGCTGAACCTCAACTGAATGGCTTAT	2152
QY	247	ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle	266
DB	2153	TTTGTGGTTGGTATATTTTGTCTATTATAACGGAGGCTGCAACACGACCATTTTCAATA	2212
QY	267	IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis	285
DB	2213	ATATTTTCAAGGATTTATAGGATCTTTACCCGAGATGAGGATCTCTGAAACAAAPAGCAG	2272
QY	286	AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr	305
DB	2273	AATAGTAACATGTTTCTGTATTGTTCTAGTCTTGGAAATTAITTTTATTATACATTT	2332
QY	306	PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis	325
DB	2333	TTCTCTCCAGGCTTCACATTTGGCAAGCTGGGAGATCTCTCACTAAGGGCTTCGATAC	2392
QY	326	LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer	345
DB	2393	ATGGTTTTTCAGATCCATGCTGAGACAGGATGTCTAGCTGTTTGTAGTCCCTTAAACACC	2452
QY	346	ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValaThrGly	365
DB	2453	ACTGAGCATTTGACAAACAGGCTTCCCAATGATGGCTCAAGTTTAAAGGGCTTAGGT	2512
QY	366	SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer	385
DB	2513	TCCAGCTTGTGTCTCATTTACCAGAAATATAGCAATCTTGGACAGGCATTAATATACC	2572
QY	386	PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal	405
DB	2573	TTAATCTATGTTGGCAATTAACACTTTACTTTAGCAATTTGTACCCTCATITCAATA	2632
QY	406	ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu	425
DB	2633	CGAGGAGTTGTTGAAATGAAATGTTGTGCAACAGCACTGAAAGATAAGAGAGGCTTA	2692
QY	426	LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu	445
DB	2693	GAAAGAGCTGGGAAGATTGCTACAGAGCCATCGAAACTTCGGAACCTGTGTTCTTTG	2752
QY	446	ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn	465
DB	2753	ACTCGGAGCAGAAAGTTTGAATACATGATGACACAGAGTTTGCAAGTACCACATACAAAAC	2812
QY	466	ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr	485
DB	2813	TCTTTGAGGAAGACACATCTTCGGGTCTCAATTTCTATCACCAGGCATGATGAT	2872
QY	486	PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr	505
DB	2873	TTTTCTATGCTGGCTGTTCCGGTTTGGTGGCTACTTTGGTGGCAATGAGTTTCATGAC	2932
QY	506	ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys	525
DB	2933	TTTCAGGATGTTCTTTGTTGTTCTCAGCTATTGTCTTTGTTGTCATGCGCAGTGGGCGAG	2992
QY	526	ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla	545
DB	2993	GTCACTTCACTTGTCTGCTGATATGCCAAAGCCAAAGTATCAGCAGCCACGCTCATG	3052

Qy	546	LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr	565
Db	3053	ATCATTGAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAGCGCTCAAGCAAAATACG	3112
Qy	566	CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal	585
Db	3113	TTGGAAGGAAATGTGACATTTTAAATGAGGTGCTGTTCAACTATCCACATCGACAGACATC	3172
Qy	586	PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly	605
Db	3173	CCCGTCTCCAGGGCTGAGCTTCGAGTGAAGAGGCCAGACGCTGCGCTCTAGGT	3232
Qy	606	SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal	625
Db	3233	AGCAGTGGCTGTGGGAAGACGACAGTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG	3292
Qy	626	GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg	645
Db	3293	GCTGGTTCAGTCTAATTGATGCGCAAGAGATATAAGCACTGAATGTCCAGTGGCTCCGA	3352
Qy	646	SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn	665
Db	3353	GCACCTGGGCGATGCTCTCAGAGGCCATCCTGTTGACTGCAGCATTCGCCAGAAC	3412
Qy	666	IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn	685
Db	3413	ATTGCTATGGAGACACACAGCCGGTCTGATCACATGAAGAGATTATGCAGCGACCAAG	3472
Qy	686	AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly	705
Db	3473	GAGGCCAACATACACCACTTCATCAGACACTCCCTCGAGAAATACAAACACAGATAGGA	3532
Qy	706	LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu	725
Db	3533	GACAAGGAACCACTCTCTGTGGTGGCAGAAACAGCGCATTTGCCATGCTCGCGCTCT	3592
Qy	726	GlnGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer	745
Db	3593	GTTAGACAGCCTCATATTTTCTTTGGATGAAGCTACATCAGCTCTGGATACAGAACT	3652
Qy	746	GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal	765
Db	3653	GAAGAAGTTGTCCAAGAGCCCTGGCAAAAGCCAGAGAAGCGCCACCTGTGATC	3712
Qy	766	ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnLys	785
Db	3713	GCCCAACCGCTTGTCCACCATCCAGAAATGAGATTTAATAGTGTGTTTCAGATGGCAAA	3772
Qy	786	IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu	805
Db	3773	GTCAGGAGCATGGCACACATCAACAGCTGCTGGCCAGAAAGGCATCTATTTTCCATG	3832
Qy	806	ValAsnAlaGlnSer	810
Db	3833	GTCAGTGTCCAGCT	3847
RESULT	8		
AAF86127			
ID	AAF86127	standard; cDNA; 4186 BP.	
XX			
AC	AAF86127;		
XX			
DT	25-JUN-2001	(first entry)	
XX			
DE		Cynomologous monkey P-glycoprotein cDNA.	
XX			
KW		Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;	
KW		efflux pump; ss.	
XX			
OS		Macaca fascicularis.	
XX			
Key		Location/Qualifiers	

FT	CDS	100..3942	
FT		/*tag= a	
FT		/product= "PGP"	
XX		/note= "P-glycoprotein"	
PN	WO200123565-A1.		
XX			
PD	05-APR-2001.		
XX			
PF	28-SEP-2000; 2000WO-US26592.		
XX			
PR	28-SEP-1999; 99US-0156921.		
PR	12-OCT-1999; 99US-0158818.		
XX			
PA	(GENT-) GENTEST CORP.		
XX			
PI	Stocker PJ, Steimel-Crespi DT, Creepi CL;		
XX			
DR	WPI; 2001-316136/33.		
DR	P-PSDB; AAB81064.		
XX			
PT	Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting bioavailability of compound and increasing PGP transporter activity in cell		
XX			
PS	Example 1; Page 51-57; 84pp; English.		
XX			
CC	This invention relates to a polynucleotide sequence encoding a cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the PGP protein. PGP, also known as multidrug transporter, MDR1 is a member of the ABC transporter superfamily. The enzyme serves as an efflux pump exporting small molecules across the cell membrane. The invention includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence and protein, and also that of an allelic variant. The PGP polynucleotide sequence is useful for increasing PGP transporter activity in a cell. Antisense sequences of the cDNA are useful for inhibiting PGP transport activity in a mammalian cell. They may also be used for increasing the bioavailability of a drug. The present sequence represents cDNA encoding cynomologous monkey P-glycoprotein.		
XX			
SQ	Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 Other;		
Alignment Scores:			
Pred. No.:	2,04e-226	Length:	4186
Score:	2441.00	Matches:	465
Percent Similarity:	76.73%	Conservative:	168
Best Local Similarity:	56.36%	Mismatches:	174
Query Match:	59.84%	Indels:	18
DB:	22	Gaps:	5
US-09-873-409-2 (1-812) x AAF86127 (1-4186)			
Qy	2	ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal	21
Db	1456	GTTGATGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTG	1515
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1516	GTCAGTCAGGAACCTGTATTTGTTCCACACGATAGTGTGAAACATTCGCTATGTGCT	1575
Qy	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
Db	1576	GAAGATGTCACCATGATGAGATTGAGAAAGCTGTCAAGAGCAATGCTATGACTTT	1635
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyAlaGlnMetSer	81
Db	1636	ATCATGAAACTCCCTCAGAAATTTGCACCCCTGTTGGAGAGAGAGGGCCCGCTGAGT	1695
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1696	GGTGGGCAGAGACAGAGGATGCCATTGCAGTGCCTTGGTTCGCAACCCCAAGATCCTC	1755



Db 3910 GTCAGTGTCCAGGCT 3924

RESULT 9

AAF86128

ID AAF86128 standard; cDNA; 4195 BP.

XX

AC AAF86128;

XX

DT 25-JUN-2001 (first entry)

XX

XX Cynomologous monkey P-glycoprotein cDNA variant 1.

DE

XX Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL1;

KW efflux pump; ss.

XX

OS Macaca fascicularis.

XX

PH Key Location/Qualifiers

PH CDS 100..3951

FT FT /\*tag= a

FT FT /product= "PGP"

FT FT /note= "P-glycoprotein"

FT FT 376..384

FT FT /\*tag= b

FT FT /note= "Insertion of 9 nucleotides relative to PGP

FT FT allelic variant AAF86127"

PN WO200123565-A1.

XX

PD 05-APR-2001.

XX

XX 28-SEP-2000; 2000WO-US26592.

PF

XX 28-SEP-1999; 99US-0156921.

PR

XX 12-OCT-1999; 99US-0158816.

PR

XX (GENT-) GENTEST CORP.

PA

XX

XX Stocker PJ, Steimel-Crespi DT, Crespi CL;

PI

XX

DR WPI; 2001-316136/33.

DR

XX P-PSDB; AAB81065.

XX

PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein

PT (PGP) and homologous PGP polypeptides are useful for predicting

PT bioavailability of compound and increasing PGP transporter activity in

PT cell -

XX

XX Example 1; Page 59-65; 84pp; English.

XX

XX This invention relates to a polynucleotide sequence encoding a

CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the

CC PGP protein. PGP, also known as multidrug transporter, MDRL1 is a member

CC of the ABC transporter superfamily. The enzyme serves as an efflux pump

CC exporting small molecules across the cell membrane. The invention

CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence

CC and protein, and also that of an allelic variant. The PGP polynucleotide

CC sequence is useful for increasing PGP transporter activity in a cell.

CC Antisense sequences of the cDNA are useful for inhibiting PGP transport

CC activity in a mammalian cell. They may also be used for increasing the

CC bioavailability of a drug. The present sequence represents cDNA encoding

CC cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide

CC insert compared to the PGP allelic variant given in AAF86127.

XX

SQ Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:

Pred. No.: 2,05e-236 Length: 4195

Score: 2441.00 Matches: 465

Percent Similarity: 76.73% Conservative: 168

Best Local Similarity: 56.36% Mismatches: 174

Query Match: 59.84% Indels: 18

DB: 22 Gaps: 5

Db 2479 ATGGTTTCCGATCCATGCTCAGACAGGATGTGAGCTGTTGTATGATCACCCCTAAACACACC 2538  
 Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365  
 Db 2539 ACTGGAGCATTCGATACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGGT 2598  
 Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385  
 Db 2599 TCCAGCCTTGCTATATATACCCAGAAATATAGCAAAATCTTGGGACAGAAATAATATATCC 2658  
 Qy 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405  
 Db 2659 TTAATCTATGTTGGCACTGACACTGTACTCTTACCAATGTACCCATCATGTGCAATA 2718  
 Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425  
 Db 2719 GCAGGAGTTGTTGAATGAATGTTGCTGGACAGCACTGAAAGATAAGAAAGACTA 2778  
 Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445  
 Db 2779 GAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGGAACCTGTGTTCTTTG 2838  
 Qy 446 ThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsn 465  
 Db 2839 ACTCAGAGCAGCAAGTTTGAACATATGATGATCAGAGTTTGACAGTACCATACAGAAAC 2898  
 Qy 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485  
 Db 2899 TCTTTGAGNAAGCACACATCTTTGGAATCACGTTTCTTCAGCGAGGCAATGATGAT 2958  
 Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArgMetThr 505  
 Db 2959 TTTTCTATGCTGGATGTTTCCGTTTGGAGCCTACTTGGTGACATATGCTCATGAGC 3018  
 Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525  
 Db 3019 TTTGAGATGCTCTGTAGTATTTTTCAGCTGTGTGTTGGTCCATGGCCGTGGGGCAA 3078  
 Qy 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyValAlaHisLeuPheAla 545  
 Db 3079 GTCAAGTTTCATTTGCTCTGACTATGCAAGCAAGCAAGTATCAGCAAGCCCATCATG 3138  
 Qy 546 LeuLeuGluLysLysProAlaIleAspSerArgSerGlnGluGlyLysLysProAspThr 565  
 Db 3139 ATCATTTGAAAAACCCCTTTGATTGACGTACAGCACAGAAAGCCCTAAAGCCGACACA 3198  
 Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585  
 Db 3199 TTGGAAGGAATATGCATTTAATGAAGTTGTATTCAACTATCCACCCGACTGGACATC 3258  
 Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605  
 Db 3259 CNAAGTCTCAGGGGCTGAGCTTGAAGTGAAGAGGGCCAGACGCTGGCCCTGGTGGGC 3318  
 Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625  
 Db 3319 AGCAGTGGCTGTGGGAAGACACGGTGGTCCAGCTCTCGAGCGGTTCTATGACCCCTTG 3378  
 Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArg 645  
 Db 3379 GCGGGGAAAGTGTCTGTACGGCAAGAAATAAAGCAACTGTAATGTTGAGTGGCTCCGA 3438  
 Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665  
 Db 3439 GCACACCTGGGCATCGTGTCCAGGAGGCCATCTGTTTACGTGACGATTAAGTGAAGAC 3498  
 Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685  
 Db 3499 ATTGCCTATGGAGACACACAGCCGGTGTGTACAGAGAGAGATCGTGGAGGGCAGCAAG 3558  
 Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705  
 Db 3559 GAGGCCAATATACACGCCCTTCATCGAGTCACTGCCCTTAATAATATAGCACAGATGAG 3618

Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaIleAlaIleAlaIle 725  
 Db 3619 GACAAAGGAACCTCAGCTCTCTGGTGGCCAGAAACACGATTTGCCATAGCTGTCGCCCTT 3678  
 Qy 726 LeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745  
 Db 3679 GTTAGCAGACCTCATATTGCTTTGGATGAAGCCACATCAGCTCTGGATACAGAAAGT 3738  
 Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765  
 Db 3739 GAAAGGTTGTCAGAAAGCCCTGGACAAAGCCAGAGAGCCGACCTGCTGATTTGATT 3798  
 Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785  
 Db 3799 GCTCACCGCTGCTCCACCATCCAGAAATCAGACTTAATAGTGGTGTTCAGAAATGGCAGA 3858  
 Qy 786 IleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeu 805  
 Db 3859 GTCAAGGAGCAGGACACATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATG 3918  
 Qy 806 ValAsnAlaGlnSer 810  
 Db 3919 GTCAAGTGTCCAGGCT 3933  
 RESULT 10  
 ID AAD03506  
 XX AAD03506 standard; cDNA; 4279 BP.  
 AC AAD03506;  
 XX 13-JUN-2001 (first entry)  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.  
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.  
 OS Canis familiaris.  
 PH Key Location/Qualifiers  
 FT CDS 17..3862  
 FT /\*tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype D) protein"  
 FT replace (91, T)  
 FT /\*tag= b  
 FT replace (607, C)  
 FT /\*tag= c  
 FT replace (1001, T)  
 FT /\*tag= c  
 FT replace (3458, A)  
 FT /\*tag= c  
 XX WO200123540-A2.  
 PN 05-APR-2001.  
 PD 28-SEP-2000; 2000WO-US26767.  
 PF 28-SEP-1999; 99US-0156510.  
 PR (GENT-) GENTEST CORP.  
 PA Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
 PI WPI; 2001-235373/24.  
 XX P-PSDB; AAE00310.  
 DR New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX



PS Claim 9; Page 102-107; 111pp; English.

CC The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity, by  
CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump  
CC exporting small molecules across the cell membrane. This enzyme  
CC is a member of the ABC transporter family.

XX SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

#### Alignment Scores:

Pred. No.:	2.11e-226	Length:	4279
Score:	2441.00	Matches:	469
Percent Similarity:	76.24%	Conservative:	160
Best Local Similarity:	56.85%	Mismatches:	178
Query Match:	59.84%	Indels:	18
DB:	22	Gaps:	4

US-09-873-409-2 (1-812) x AAD03506 (1-4279)

Qy	2	ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal	21
Db	1379	ATTGATGACAGACATTAAGACCATAAATGTAAAGCATCTTCGGGAATTAATCTGGTGTG	1438
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1439	GTGAGTCAGGAGCCTGTGTGTTGGCCACACGATAGCTGAAACATTCGCTATGCGCCG	1498
Qy	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
Db	1499	GAATAATGTCACCATGGATGAGATGAGAAAGCTGTAAAGAAAGCCAAATGCCCTATGATTT	1558
Qy	62	IleMetGluPheProAsnLysPheAenThrLeuValGlyGlyValAlaGlnMetSer	81
Db	1559	ATCATGAACACTACCTAATAATTTGACACTGTGTTGGAGAGAGGGGCCACGCTGAGT	1618
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1619	GCTGGACAGAAACACAGAAATGCCATTGCTCGGGCCCTGTTCCGAACCCCAAGATTCTT	1678
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1679	CTGTGGATGAGCAACGTCAGCTCTGGACACTGGAAGTGAAGCAGTGTTCAGGTGGCC	1738
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIle	141
Db	1739	CTGGATAGCCGAGAAAGCCGACTACATTTGTATAGCTCATCTTTGTGTCTACAGTT	1798
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1799	CGTAATGCCGATGTCATTGCTGTTTTGTATGATGGAGTCATTGTGGAGAAAGAAATCAT	1858
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys	181
Db	1859	GATGAACACTCATGAAGAGAGGCGCATTTACTTCAAACTTGTCCAAATGCAG---ACAAGA	1915
Qy	182	LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu	201
Db	1916	GGAAATGAATTAAGATTGAATAATGCCATCTGTGTAATCCAAAGTCAAAAGTATGCTCTG	1975
Qy	202	ProLeuHisSerValLysSerIleLysSerAspPheIle-----	214

Db	1976	GAATG---TCTCCAAAAGATTTCAGGGTCCAGTTTATAAATAAAGAGATCAACTCGCAGG	2032
Qy	215	-----AspLysAlaGluGluSerThrGlnSerLysGluIle	226
Db	2033	AGTATACATGCACCACCAAGGCCACAGACAGAAAGCTTGGTACAAAGAGAGACTTGAATGAG	2092
Qy	227	SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGluTyrProPhe	246
Db	2093	AATGTACCTCCAGTTTCTCTCTGGAGATTCTGAAGCTGAACCTCAACTGAATGGCTTAT	2152
Qy	247	ValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIle	266
Db	2153	TTTGTGGTGTGTATATTTTGTCTATTATAACCGAGGCGCTGCAACACGACATTTTCAATA	2212
Qy	267	IlePheAlaLysIleIleThrMetPhe---GlyAenAsnAspLysThrThrLeuLysHis	285
Db	2213	ATATTTTCAAGATTAATAGGATCTTTACCCGAGATGAGGATCTCTGAACAAACACACAG	2272
Qy	286	AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr	305
Db	2273	AATAGTAACATGTTTCTGTATTGTTCTAGTCCTTGAATTAATTTCTTTATTATCAATTT	2332
Qy	306	PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis	325
Db	2333	TTCTCTCCAGGGCTTCACATTTTGGCAAGCTGGGGAGATCTCTACATAAGCGGCTTCGATAC	2392
Qy	336	LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer	345
Db	2393	ATGCTTTTCAGATCCATGCTGAGACAGAGTGTACGCTGGTTGTATGACCCCTAAACACCC	2452
Qy	346	ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly	365
Db	2453	ACTGGAGCATTTGACAAACACCGCTTGCCCAATGATGGCGCTCAAGTTAAAGGGCTTAGGT	2512
Qy	366	SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer	385
Db	2513	TCCAGGCTTGCTGTCTATTACCCAGAAATATAGCAATCTTGGGACAGGCATTTATTATCC	2572
Qy	386	PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal	405
Db	2573	TTAATCTATGTTGGCAATTAACATTTTCTTATGACAAATGTACCCATCATATTGCNATA	2632
Qy	406	ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu	425
Db	2633	GCAGGAGTGTGTTGAATGAAATGAAATGTTGCTGGCAAGCAGCTGAAAGATAAGAGAGCTA	2692
Qy	426	LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu	445
Db	2693	GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACCTGTGTTCTTTG	2752
Qy	446	ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn	465
Db	2753	ACTCGGAGCAGAGATTTGAATACATGATGACACAGAGTTTGCAAGTACCATACAGAAC	2812
Qy	466	ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr	485
Db	2813	TCTTTGAGGAAGACACATCTTCGGGGTCTCATTTTCTTCTACCCAGGCAATGATGAT	2872
Qy	486	PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr	505
Db	2873	TTTTCTCTATGCTGGGTGTTTCGGTTCGCTACTTGTGGTGGCAAAATGAGTTCATGAAC	2932
Qy	506	ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLys	525
Db	2933	TTTCAGGATGTTCTTTGTGTTATCTCAGCTATGTTCTTTGGTGGCCATGGCAGTGGGGCAG	2992
Qy	526	ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla	545
Db	2993	GTCAGTTTCATTTGCTCTCTGACTATGCGAAAGCCAAAGTATCAGCAGGCCACGCTCATG	3052
Qy	546	LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr	565

```
Db 3053 ATCATTTGAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAGGCTCAAGCCAAATACG 3112
Qy 566 CyeGluGlyAsnLeuGluPheAArgGluValSerPhePheTyrProCysArgProAspVal 585
Db 3113 TTGGAAGGAAATGTGACATTTAATGAGTCTGTGTTCAACTATCCACTGACACAGACATC 3172
Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 3173 CCCGTGCTCCAGGGGCTGAGCCTCGAGTGAAGAGGCCAGAGCGCTGGCCCTCGTAGGT 3232
Qy 606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db 3233 AGCAGTGGCTGTGGGAAGACACAGTTGTTGAGCTCTAGAGCGCTCTATGACCCCTTG 3292
Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db 3293 GCTGGTTTCAGTGCTAATTGATGTCGAAGAGATAAAGCACCTGATGTCCAGTGGCTCCGA 3352
Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db 3353 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCTGTTGACTGCAGCATTTGCCGAGAAC 3412
Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db 3413 ATTGCCTATGGAGACACACAGCCGGTCTATCATGTAAGAGATTGTGCAGGAGCCAAAG 3472
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db 3473 GAGGCCAACATACACACTTTCATCGACACTCCCTGAGAAATACAAACACAGGATAGGA 3532
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db 3533 GACAAAGGAACCCAGGTCTCTGTGGGCCAGAAACAGCGCATTCCTCATAGTCGCGCTCT 3592
Qy 726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db 3593 GTTAGACAGCTCATATTTTGCTTTTGGATGAGCTACATCAGCTCTGGATACAGAAAGT 3652
Qy 746 GluLysValValGlnIleHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db 3653 GAAAGAGTTTCCAAAGAGCCCTGGCAAGCCAGAGAGGCCGACCTGCTGATTTGATC 3712
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspIleValValLeuHisAsnGlyLys 785
Db 3713 GCCCAGCGCTGTGCCACCATTCAGATGCGAGATTTAATAGTGTGTTTCAGAAATGCCAAA 3772
Qy 786 IleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Db 3773 GTCAGGAGCATGGCACATCATCAACAGCTCTGGCCAGAAAGGCATCTATTTTTCATG 3832
Qy 806 ValAsnAlaGlnSer 810
Db 3833 GTCAGTGTCCAGGCT 3847
RESULT 11
AAZ52047
ID AAZ52047 standard; cDNA; 4369 BP.
XX
AC AAZ52047;
XX
DT 18-JUL-2000 (first entry)
XX
DE Rat multidrug resistance protein 1a cDNA.
XX
KW Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;
KW drug formulation; formulation excipient; compound design; inflammation;
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW central nervous system disorder; auto-immune disease; kidney disease; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT CDS 352..4170
```

```
FT
FT
XX
PN WO200015650-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1999; 99WO-US20770.
XX
PR 17-SEP-1998; 98US-0156800.
PR 09-DEC-1998; 98US-0208809.
XX
PA (SMK ) SMITHKLINE BEECHAM CORP.
XX
PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
XX
DR WPI: 2000-271372/23.
DR P-PSDB; AAY70596.
XX
Isolated rat mdrla polynucleotides and polypeptides, useful in assays
to provide information on drug formulation, selection of formulation
excipients and compound design .
XX
Claim 2; Page 23-24; 33pp; English.
XX
The present cDNA sequence encodes rat multidrug resistance
protein 1a (mdrla). This protein is a member of the multi-specific drug
transporters family. Mdr1a is used in assays to provide information on
drug formulation, selection of formulation excipients and compound
design. They are used in cell based, membrane based, binding or other
assays to provide information that may enhance drug formulation. This
invention further relates to the generation of in vivo and in vitro
comparison data to predict oral absorption and pharmacokinetics. This
enables the selection of drugs with optimal pharmacokinetics, i.e. good
oral bioavailability, brain penetration, plasma half life, and minimum
drug interaction. Transgenic and knock-out animals created using rat
mdrla provides an insight into treating and preventing human diseases
including cancer, inflammation, cardiovascular disease, central nervous
system disorders, auto-immune and kidney disease.
XX
SQ Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:
Pred. No.: 2.18e-226 Length: 4369
Score: 2441.00 Matches: 470
Percent Similarity: 77.04% Conservative: 164
Best Local Similarity: 57.11% Mismatches: 175
Query Match: 59.84% Indels: 14
DB: 21 Gaps: 5

US-09-873-409-2 (1-812) x AAZ52047 (1-4369)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGTATCTCGGGAAATCATTTGGGGTG 1743
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1744 GTGAGTCAGAACCCGCTGCTGTTGCCACCACCAATTCGCCGAAACATTCGCTATGSCCGA 1803
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1804 GAAAACGTCACCATGATGAGTAAAGCTGTCAAGAAAGCCCAATGCCCTATGATGTTTC 1863
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSer 81
Db 1864 ATCATGAAACTCCCAAAATTTGACACCTGTTGTGGAGAGGGCGGCGAGCTGAGT 1923
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1924 GGGGACAGAAACACAGAGGATGCCATTCGCCGGGCTGTCGCCCAACCCCAAGATCCTT 1983
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
```

Db 1984 TTTGGTGGATGAGCCACGTCACCCCTTGGACACAGAAAGCCGCTGGTTCAGGCCGCT 2043  
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 2044 CTGGATAGGCTAGAGAGCCCGGACCACCATTTGTATAGCTACCGCTTGTCTACAGTT 2103  
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
Db 2104 CGCAATGCTGACGTCATTCCTGGTGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2163  
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
Db 2164 GATGAGCTCATGAGAGAGAAAGAAATTTACTTCAAACTTGTCTGACTCAGACAGCAGA 2223  
Qy 179 ---AspIleLysAlaAspGluGlnMetGluSer-----MetThr 191  
Db 2224 AATGAATTTGAATTAGGAATAGAGCTTGTGTAATCTAAGACGGGAATGTAATGTGGAC 2283  
Qy 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209  
Db 2284 ATGTCTTCAAAAGATTCCGGATCCAGTCTTAATAAGAAAGAGATCAACTCGCAAAAGCATC 2343  
Qy 210 LysSerAppPheIleAspLysAlaGluLysThr---GlnSerLysGluIleSerLeu 228  
Db 2344 CGTGGGCCACATGATCAAGACGGGGAACCTTAGCACCAAGAGGGCTCTGGATGACGACGTA 2403  
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248  
Db 2404 CTCACAGCTTCTTTGGCGGATCCTGAAGTTGAATTCACATGAATGGCTTATTTTTGG 2463  
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
Db 2464 GTTGGTGTATTTTGTGCCATAATAATAGGAGCTTGCAACACGATCTCTCCATAATATT 2523  
Qy 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrLeu---LysHisAspAla 287  
Db 2524 TCAAAAGTTGTAGGGGTTTTTACAAAAATATGACACCCCTTGAAATCCAGCGGACAGACAGC 2583  
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
Db 2584 AACTTTTCTTTTATGTCTCTGATCCTTGGATCATCTTTTCAATACGTTTTTCTT 2643  
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
Db 2644 CAGGGCTTCACATTTGGCAAGCTGGAGAGATCCTCACCAAGGACCTCCGATACATGTC 2703  
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
Db 2704 TTTCAATCCATGCTGAGACAGACATAAGCTGTTTGTATGATGACCTTAAACACACACAGGA 2763  
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValThrGlySerArg 367  
Db 2764 GCGCTGACCACCGGCTTGCCAAATGACGCTGCTCAAGTGAAGGGGCTACAGGGTCTAGG 2823  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 2824 CTTCGCTTATTTACCAGAACATAGCAATCTTTGGACAGGCAATCATATATCCCTGATC 2883  
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2884 TACGGCTGGCAATTGACACTTTTACTCTCTAGCAATTTGTTCCCATCTATCTCTATAGCAGA 2943  
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHis 427  
Db 2944 GTGGTTGAAATGAAATGTTGTCTGGCAACGCGCTGAAAGATGAAGAAAGCAACTAGAAGT 3003  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
Db 3004 TCTGGAGAGATGCTACAGAGCAATTTGAACCTTTCCGACACTGCTCTCTTTGACTCGG 3063  
Qy 448 GlyLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467

Db 3064 GAGCAGAAAGTTTGAACCTATGTATGCCAGAGCTTGCAGATACCATACAGAAATCCTTTG 3123  
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 3124 AAGAAAGCGCAGCTTTTGGGATCACTTCTCCCTTACCAGGCCATGATGATTCTTCC 3183  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 3184 TATGCTGCTGTTTCCGCTTTCATGCTTACTTGTGTGGCAGCAAACTCATGACATTTGAA 3243  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
Db 3244 AATGTTCTGTAGTATCTCAGCTATTGCTTTGTGTCATGGCAGTGGGCGCGGTCACT 3303  
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
Db 3304 TCATTCTGCTCTGACTACGCGAAAGCCAAAGTCTCGGCATCCACATCATCAGGATCAT 3363  
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
Db 3364 GAGAAATCCCTCTGAGATTACAGCTACAGCAGCGAGGGCTTGAAGCCTAATATGTTGAA 3423  
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
Db 3424 GGAATGTGAATTTAATGGAGTCATGTTCAACTATCCACCCGACCCCAACATCCAGTG 3483  
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgLysThrValAlaPheValGlySerSer 607  
Db 3484 CTTACGGGCTGAGCTAGAGGTGAAGAAAGGCGAGACGCTGGCCTCGTGGCGAGCAGT 3543  
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627  
Db 3544 GGTCTCGGGAAGAGTACAGTGGTCTGAGCTGCTTGTAGCGCTTATGACCCCATCGCGGA 3603  
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGln 647  
Db 3604 ACAGTGTCTTAGTAGGCAAAAGAAATCAAGCAACTCAATGTCCAGTGGCTCCGCGCCAC 3663  
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667  
Db 3664 CTGGGCATTTGTCTCCAGAGCCCATCTCTTTGTACTGAGCATCGCCGAGAACATTGCG 3723  
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687  
Db 3724 TACGAGACACACAGCCGCTGCTGTCTATAAGAGAGATCGTGAAGCAGCCAGGAGGCC 3783  
Qy 688 AsnIleHisSerPheIleGluClyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707  
Db 3784 AACATCCACCACTTCATCGACTCACTGCTGAGAAATACACACAGAGTGGGAGCAAAA 3843  
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727  
Db 3844 GGGACTCAGCTGTGCGGGCGGAGAGAGCGCATCGGCATCGCGCGGCCCTCTGTGAGA 3903  
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsnAspSerGluLys 747  
Db 3904 CAGCCTCATCTTACTTCTGGATGAAGGACATCAGCTCTGATACGAGAGTGAAGAAG 3963  
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
Db 3964 GTCGTCCAGGAAGCGCTGCAAAAGCCAGGAGGCGCACCTGATTTGTGATCGCGCAC 4023  
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLys 787  
Db 4024 GCGCTGTCCACCATCCAGAACGCACTTGATCGTGGTGAATTCAGAACGCCCGCGGTCAAG 4083  
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807  
Db 4084 GAGCAGCGCACCCACCACGAGCTGTGCGCCCAAGGAGCATCTATTTCTCGATGTGCTCA 4143  
Qy 808 AlaGlnSer 810  
Db 4144 GTGCAGGCT 4152

## RESULT 12

AAZ52048

ID AAZ52048 standard; cDNA; 4425 BP.

XX AC AAZ52048;

XX DT 18-JUL-2000 (first entry)

DE DE Rat multidrug resistance protein 1a cDNA derived from EST sequences.

XX Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;

KW drug formulation; formulation excipient; compound design; inflammation;

KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;

KW central nervous system disorder; auto-immune disease; kidney disease;

KW EST; expressed sequence tag; ss.

XX OS Rattus rattus.

XX FH Key

FT CDS 352..4170

FT Location/Qualifiers

FT /\*tag= a

FT /product= "Rat multidrug resistance protein"

XX PN W0200015650-A1.

XX PD 23-MAR-2000.

XX PF 10-SEP-1999; 99NO-US20770.

XX PR 17-SEP-1998; 98US-0156800.

XX PR 09-DEC-1998; 98US-0208809.

XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;

XX DR WPI; 2000-271372/23.

XX DR P-PSDB; AAY70597.

XX PT Isolated rat mdr1a polynucleotides and polypeptides, useful in assays

XX PT to provide information on drug formulation, selection of formulation

XX PT excipients and compound design -

XX PS Claim 11; Page 27-29; 33pp; English.

CC The present cDNA sequence encodes rat multidrug resistance protein 1a

CC (mdr1a). This cDNA is derived from EST (expressed sequence tag)

CC sequences. Mdr1a is used in assays to provide information on

CC drug formulation, selection of formulation excipients and compound

CC design. They are used in cell based, membrane based, binding or other

CC assays to provide information that may enhance drug formulation. This

CC invention further relates to the generation of in vivo and in vitro

CC comparison data to predict oral absorption and pharmacokinetics. This

CC enables the selection of drugs with optimal pharmacokinetics, i.e. good

CC oral bioavailability, brain penetration, plasma half life, and minimum

CC drug interaction. Transgenic and knock-out animals created using rat

CC mdr1a provides an insight into treating and preventing human diseases

CC including cancer, inflammation, cardiovascular disease, central nervous

CC system disorders, auto-immune and kidney disease.

XX SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

## Alignment Scores:

Pred. No.: 2,22e-226 Length: 4425

Score: 2441.00 Matches: 470

Percent Similarity: 77.04% Conservative: 164

Best Local Similarity: 57.11% Mismatches: 175

Query Match: 59.84% Indels: 14

DB: 21 Gaps: 5

US-09-873-409-2 (1-812) x AAZ52048 (1-4425)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21

Db 1684 ATCGACGGACAGGACATCAGGACCAATGTGAGGTATCTGCGGAAATCAATGGGGTG 1743

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41

Db 1744 GTGAGTCAGGAACCCGCTGCTGTTCACCAATATGCCGAAACAAATTCGTATGGCCGA 1803

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61

Db 1804 GAAACGTCACCATGATGATAGAGAAAGCTGTCAAGAGCCCAATGCCTATGATTTC 1863

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81

Db 1864 ATCATGAAACTGCCCCACAAATTTGACACCTTGTTGGTGAGAGAGGGCGCAGCTGAGT 1923

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

Db 1924 GGGGACAGAAAACAGAGGATCGCCATTGCCGGGCCCTGTCGCAACCCCAAGATCCTT 1983

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121

Db 1984 TTGTTGATGAGGCCACGTCAGCTTGGAACAGAAAGCAAGCCGTGGTTACGGCCGCT 2043

QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141

Db 2044 CTGGATAAGGCTAGAGAAGCGGACCACTTGTGATAGCTCACCGCTTGTCTACAGTT 2103

QY 142 ArgSerAlaAspLeuValThrLeuLysAspGlyMetLeuAlaGlyLysGlyAlaHis 161

Db 2104 CGCAATGCTCAUGTCTGCTGCTGTTTGTATGGTGGTGTGTCATTGTGGAGCAAGCAATCAT 2163

QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178

Db 2164 GATGAGCTCATGAGAGAGAAAGAAATTACTTCAAACTTGTCTAGTACTCAGACAGCAGGA 2223

QY 179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----MetThr 191

Db 2224 AATGAAATTTGAATAGGAATGAAGCTTGTGAATCTAAAGAUUGAATTCATAATGGGAC 2283

QY 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209

Db 2284 ATGTCTCTCAAAAGATTCRGGATCCAGTCTTAATAAGAAAGAAATCAACTCGCAAAACATC 2343

QY 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 238

Db 2344 CGTGGCCACATGATCAACACGGGAACTTAGCACCAAGAGGCTCTGGATGACGACGTA 2403

QY 229 ProGluValSerLeuLysLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248

Db 2404 CCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCTTATTTTGTG 2463

QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268

Db 2464 GTTGGTGTATTTTGTGCCATAATAATGAGGCTTGCAACACGACATTCCTCCATAATATT 2523

QY 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrLeu---LysHisAspAla 287

Db 2524 TCAAAAGTTGTAGGGGTTTTTCAAAAATGACACCCCTGAAATCCAGCGGAGAAACAGC 2583

QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307

Db 2584 AACTGTGTTTCTTTATTGTTCTGATCCTTGGGATCATCTCTTCATTAGGTTTTTCTT 2643

QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327

Db 2644 CAGGGCTTCACATTTGGCAAGCTGGAGAGATCCTCACAAGCGACTCCGACATCATGTC 2703

QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTropPheAspGlyLysGluAsnSerThrGly 347

Db 2704 TTCAAAATCCATGCTGAGACAGGACATAAGCTGTTGTATGATGACCTTAAACACACAGGA 2763

QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367

Db 2764 GCGCTGACCCAGCGCTTGCCTCAAGTCAAGTCAAGGAGGCTACAGGGTCTAGG 2823  
Qy ILeGlyValLeuThrGlnAsnAlaThrAenMetGlyLeuSerValIleIleSerPheIle 387  
Db 2824 CTTGCTGTATTACCCAGACATACCAATCTTGGGACAGGCATCATATATCCCTGATC 2883  
Qy TyrGlyTTPGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2884 TAGCGCTGGCAATTGACACTTTTACTCTAGCAATTTGTTCCCATCTATGCTATACAGGA 2943  
Qy MetIleGluThrAlaAlaMetThrGlyPheAlaAenLysAspLysGlnLeuLysHis 427  
Db 2944 GTGGTGGAAATGAAATGTGTCTGGCAAGCGCTGAAAGATGAAAGAACTAGAAAGT 3003  
Qy AlaGlyLysIleAlaThrGluAlaLeuGluAenIleAArgThrIleValSerLeuThrArg 447  
Db 3004 TCTGGGAGATCGCTACAGACCAATTGAAACTTTGGCACTGTCTCTCTTTGACTCGG 3063  
Qy GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisAArgAenThrSer 467  
Db 3064 GAGCAGAAGTTTGAAGTATGTATGCCAGAGCTTGCAGATACCATACAGAAATGCTTTG 3123  
Qy LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 3124 AAGAAAGCGCAGCTTTTGGGATCACTTTCTCTCCACCGGCCCATGATGTATTTCTCC 3183  
Qy TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 3184 TATGCTGCTGTTTCCGGTTTGATGCCCTACTTGGTGGGACGAGAACTCATGACATTTGAA 3243  
Qy GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
Db 3244 AATGTTCTGTAGTATTCTCAGCTATTGCTTTGGTGCATGCGAGTGGGCGAGTCACT 3303  
Qy ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
Db 3304 TCATTGCTCTCTGACTAGCGGAAAGCAAAAGTCTCGGCATCCACATCATCAGGATCAT 3363  
Qy GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
Db 3364 GAGAAATCCCTTGAGATTGACAGCTACGACGAGGAGGCTTGAAGCCTTAATATGTTGAA 3423  
Qy GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
Db 3424 GGAATGTGAATTTAATGAGTCTATGTTCACTATCCACCCGACCCCAACATCCAGTG 3483  
Qy LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607  
Db 3484 CTTACGGGCTGAGCTAGAGGTGAAGAAAGGCGACGCTGGCCCTCTGTGGGCGAGCT 3543  
Qy GlyCysGlyLysSerThrSerValGlnLeuLeuGlnAenGlyLeuTyrAspProValGlnGly 627  
Db 3544 GGCTCGGGAGAGTACAGTGTGTCAGCTGTGTGAGCTTCTATGACCCCATGGCCGGA 3603  
Qy GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrIleArgSerGln 647  
Db 3604 ACAGTGTTCATAGTGGCAAGAAATGAAGCACTCAATGTCCAGTGGTCCGCGCCAC 3663  
Qy IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAenIleAla 667  
Db 3664 CTGGGCAATTGTCTCCAGGAGCCCATCTCTTTGACTGCAGCATCGCCGAGAACTTGC 3723  
Qy TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687  
Db 3724 TACGGAGACAAACGCGGTGCTGTGTTCTAATAGGAGATCTGTGAAGCAGCCCAAGGAGCC 3783  
Qy AenIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707  
Db 3784 AACATCCACCACTTATCATGCTACCTGCTGAGAAATACAAACACAGAGTGGGAGACAA 3843  
Qy GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727

Db 3844 GGGACTACGTGTGCGGGCGGCGAGAGCAGCGCATCGCCATCGCGCGCCCTCTCTCAGA 3903  
Qy LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747  
Db 3904 CAGCCTCACATCTTACTTCTGATGAACGACATCAGCTCTGGATACGAGAGTGAAGAAG 3963  
Qy ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
Db 3964 GTCGTCCAGGAAGCGCTGGACAAAGCCAGGAAAGGCCACCTGCATTTGTGATCGCGCAC 4023  
Qy ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLys 787  
Db 4024 CCCTGTCCACATCCAGACGACATTTGATCGTGTGATTTTCAGACGCGCGAGTCAAG 4083  
Qy GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807  
Db 4084 GAGCAGCGCACCACACAGCTGTGCGCCAGAAAGGCATCTATTCTCGATGTGTCAGT 4143  
Qy 808 AlaGlnSer 810  
Db 4144 GTGCAGGCT 4152  
RESULT 13  
RAD03489  
ID AAD03489 standard; cDNA; 4317 BP.  
XX  
AC AAD03489;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Dog P-glycoprotein (PGP) cDNA #2.  
XX  
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDRI;  
KW drug bioavailability; transgenic animal; genetic model; ss.  
XX  
OS Canis familiaris.  
XX  
FH Key Location/Qualifiers  
CDS 70..3912 /\*tag= a /product= "Dog P-glycoprotein (PGP) #2"  
XX  
FN WO200123540-A2.  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-US26767.  
XX  
XX 28-SEP-1999; 99US-0156510.  
XX  
XX (GENT-) GENTEST CORP.  
XX  
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
XX WPI; 2001-235373/24.  
XX P-PSDB; AAE00304.  
XX  
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
XX for determining the bioavailability of drugs and for screening for dog  
XX PGP inhibitors -  
XX  
XX Claim 1; Page 66-72; 111pp; English.  
XX  
XX The invention relates to dog P-glycoprotein (PGP) also referred  
XX as multidrug transporter (MDRI) and nucleic acids encoding them.  
XX The invention also includes fragments and biologically functional  
XX variants of dog P-glycoprotein. PGP and their nucleic acids are  
XX useful for determining the bioavailability of drugs and for  
XX screening PGP inhibitors. They are useful for the diagnosis and  
XX treatment of conditions characterised by PGP activity, by  
XX reducing or increasing PGP activity in a cell. PGP nucleic acids  
XX are used as oligonucleotide probes. Complements of PGP nucleic  
XX acids are useful as antisense oligonucleotides, to induce a PGP

'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) cDNA. The  
 CC PGP enzyme functions as an efflux pump exporting small molecules  
 CC across the cell membrane. This enzyme is a member of the ABC  
 CC transporter family.  
 XX

SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:  
 Pred. No.: 4,18e-226 Length: 4317  
 Score: 2438.00 Matches: 467  
 Percent Similarity: 76.24% Conservative: 162  
 Best Local Similarity: 56.61% Mismatches: 178  
 Query Match: 59.77% Indels: 18  
 DB: 22 Gaps: 4

US-09-873-409-2 (1-812) x AAD03489 (1-4317)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1429 ARTGATGGACAGACATTAGGACCAATAATGTAAAGCATCTTCGGGAATATTACTGGTGTG 1488  
 Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1489 GTGAGTCAGAGGCTGTGTGTTGGCCACACGATAGCTGAAAAACATTGCTATGCGCGC 1548  
 Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 Db 1549 GAAATGTCCACCATGATGAGATTGGAAGCTGTAAAGGAAGCCATGCTTATGATTTT 1608  
 Qy 62 IleMetGlnPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1609 ATCATGAACACTACTATAAATTGACATCTGTTGGAGAGAGGGCCCGCTGAGT 1668  
 Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1669 GGTGGACAGAAACAGAGATCGCATTTGCTCGGGCCCTGTTTCCCAACCCCAAGATTCTT 1728  
 Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 1729 CTGCTGGATGAGCAACGTCAGCTCTGGACACTGGAAGTGAAGCATGGTTCAGGTGGCC 1788  
 Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
 Db 1789 CTGGATAAGCCAGAAAGCCCGGACTACCATTTGTGATAGCTCATCTGTTGTCTACAGTT 1848  
 Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
 Db 1849 CGTAATGCCGATGTCATTGCTGTTTGTATGATGATGAGTCAATTGGAGAAAGGAAATCAT 1908  
 Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181  
 Db 1909 GATGAATCATGATAAGAGAGGCGCATTTACTCAAACTTGTCTCAATGTCAG---ACAAGA 1965  
 Qy 182 LysAlaAspGluGlnMetGluSerMetTyrSerThrGluArgLysThrAsnSerLeu 201  
 Db 1966 GGAATGAATGAGTTAGTAAGAAATGCCACTGTGTAATCCAAAGTGAAGTGAATGCTTG 2025  
 Qy 202 ProLeuHisSerValLysSerIleLysSerAspPheIle----- 214  
 Db 2026 GAAATG---TCTCCAAAGATTCAGGCTCCAGTTTAAATAAAGAAAGATCAACTCGCAGG 2082  
 Qy 215 -----AspLysAlaGluSerThrGlnSerLysGluIle 226  
 Db 2083 AGTATACATGCCACACAGGCCAACAGCAAGCTGGTTACAAAGAGGACTTGAATGAG 2142  
 Qy 227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluIleTyrProPhe 246  
 Db 2143 AATGTTCCCTCAGTTTCCTCTGAGGATTTCTGAGCTGAACTCAACTGAATGCTTAT 2202  
 Qy 247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266

Db 2203 TTTGGTGGTGGTATATTTTGTGCTATTATAAACGGAGGCTCCAACCCAGCATTTTCAATA 2262  
 Qy 267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285  
 Db 2263 ATATTTTCAAGGATTTATAGGATCTTTACCCGAGATGAGGATCTCTGAAACAAACAGCAG 2322  
 Qy 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305  
 Db 2323 AATAGTAACATGTTTCTGTATTGTTCTAGTCTTGGAAATATTCTTTTATTATCAATT 2382  
 Qy 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325  
 Db 2383 TTTCTTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCTCTCACTAAGCGGCTTCGATAC 2442  
 Qy 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer 345  
 Db 2443 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGTTTGTGATGACCCCTAAAAACACC 2502  
 Qy 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365  
 Db 2503 ACTGGAGCATTTGACAAACAGGCTTGCCAAATGATGGGCTCAAGTTTAAAGGGCTATAGT 2562  
 Qy 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385  
 Db 2563 TCCAGGCTTGTGCTCATACCAGAAATATAGCAATCTTTGGGACAGGCATTATTATATCC 2622  
 Qy 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405  
 Db 2623 TTAATCTATGTTGGCAATTAACACTTTACTCTTAGCAATTGTACCCATCATTTGCAATA 2682  
 Qy 406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425  
 Db 2683 GCAGGAGTGTGTTGAAATGAAATGTTGTCGCAACAGCACTGAAAGATTAAGAAAGAGCTA 2742  
 Qy 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445  
 Db 2743 GAAGAGCTGGGAAGATTTGTCAGAACCATCGAAACTTCGAACTGTTGTTCTTTT 2802  
 Qy 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465  
 Db 2803 ACTCGGACAGCAAGTTTGAATACATGATGTCACAGAGTTTGCAGTACCATACAGAAAC 2862  
 Qy 466 ThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485  
 Db 2863 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTCTTCTATCACCAGGCAATGATGAT 2922  
 Qy 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505  
 Db 2923 TTTTCTATGCTGGCTGTTTTCGGTTCGCTACTTGTGGCAAAATGAGTTTCATGAAC 2982  
 Qy 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525  
 Db 2983 TTTCAGGATGTTCTTTGTTGTTCTCAGCTATGTTCTTTGGTGCCATGGCAGTGGGGCAG 3042  
 Qy 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545  
 Db 3043 GTCAGTTTCATTTGCTCTGACTATGCCAAAGCAAAAGTATCAGCAGCCCAACGTCATCATG 3102  
 Qy 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565  
 Db 3103 ATCATTTGAAAGAGCCCTCTGATTGACAGTACAGCCCTCAGCGGCTCAAGCCAAATACG 3162  
 Qy 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585  
 Db 3163 TTGGAAGGAATGTGATTTAATGAGGTCTGTTTCAACTATCCACTGCACAGACATC 3222  
 Qy 586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605  
 Db 3223 CCGTGTCTCCAGGGGCTGAGCCTCGAGGTGAAGAGGGCCAGACGCTGGCCCTCGTAGGT 3282  
 Qy 606 SerSerGlyCysGlyLysSerThrValGlnLeuGlnLeuGlnLeuTyrAspProVal 625



Db 3283 AGCAGTGGCTGTGGGAAGCAGCAGTGTTCAGCTCTTAGACGGCTTCTATGACCCCTTG 3342  
Qy 626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645  
Db 3343 GCTGTTTCAGTCTAATTGATGGCAAGAGATAAGCACCTGAATGTCAGTGGCTCGA 3402  
Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665  
Db 3403 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCTGTTTGAAGTTCAGGATTCAGGAGAAC 3462  
Qy 666 IleAlaTyGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685  
Db 3463 AFTGCTATGGAGACAACAGCCGGTCTGATCACATGAAGAGATTATGAGGAGCCCAAG 3522  
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyAsnThrGlnValGly 705  
Db 3523 GAGGCCACATACACCTTCATCGAGACACTCCCTGAGAAATACACACGAGTAGGA 3582  
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725  
Db 3583 GACAAGGAACCCAGCTCTCTGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3642  
Qy 726 LeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnSer 745  
Db 3643 GTTAGACAGCCCTCATATTTTGTGTTGGATGAAGCTACATCAGCTCTGGATACAGAAAT 3702  
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765  
Db 3703 GAAAGTTGTCCAGAACCCCTGGCAAGCCAGAGAGCGCGACCTGCTGATGATC 3762  
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785  
Db 3763 GCCCACCCTGTCCACCATCCAGAAATGCAGATTTAATAGTGTGTTTCAGAAATGGCAAA 3822  
Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrrPheLysLeu 805  
Db 3823 GTCAAGGAGCATGGCACATCAACAGCTGCTGGCTCAGAAAGGCGCATCTATTTTCCATG 3882  
Qy 806 ValAsnAlaGlnSer 810  
Db 3883 ATCAGTGTCCAGCT 3897  
RESULT 14  
ID ABL91687  
XX ABL91687 standard; DNA; 3840 BP.  
XX AC ABL91687;  
XX  
DT 28-MAY-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 30.  
XX  
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
KW Cytostatic; virucide; protozoacide; antibacterial; ds.  
XX  
OS Homo sapiens.  
XX  
PN DE10100586-C1.  
XX  
PD 11-APR-2002.  
XX  
PF 09-JAN-2001; 2001DE-1000586.  
XX  
PR 09-JAN-2001; 2001DE-1000586.  
XX  
PA (RIBO-) RIBOPHARMA AG.  
XX  
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
XX  
DR WPI; 2002-270454/32.  
XX  
PT Inhibiting gene expression in cells, useful for e.g. treating tumors,

PT by introducing double-stranded complementary oligoRNA having unpaired  
PT terminal bases -  
PS Claim 13; Page 28-30; 104pp; German.  
XX  
CC The invention relates to a method for inhibiting expression of a target  
CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
CC oligoribonucleotide that has a double-stranded structure consisting of at  
CC most 49 sequential nucleotide pairs, with at least part of one strand  
CC complementary with the target gene and has at least one end a  
CC single-stranded segment of 1-4 nt. The method provides  
CC oligoribonucleotides for antisense inhibition of gene expression useful  
CC e.g. for treating tumours but the oligoribonucleotides may also be  
CC directed against genes present in pathogens (e.g. Plasmodium or  
CC viruses/viroids, pathogenic on humans, animals or plants) or against  
CC cytokine, Id, developmental or prion genes. The method provides more  
CC effective inhibition of gene expression than use of known  
CC oligonucleotides, probably because the unpaired overhang increases  
CC stability and thus intracellular concentration.  
XX  
SQ Sequence 3840 BP; 1130 A; 738 C; 957 G; 1015 T; 0 other;

Alignment Scores:  
Pred. No.: 5,51e-226 Length: 3840  
Score: 2436.00 Matches: 466  
Percent Similarity: 76.79% Conservative: 166  
Best Local Similarity: 56.62% Mismatches: 177  
Query Match: 59.72% Indels: 14  
DB: 24 Gaps: 4

US-09-873-409-2 (1-812) x ABL91687 (1-3840)

Qy 2 ValAspGluAsnAspIleAargAlaLeuAsnValArgHisTyArgAspHisIleGlyVal 21  
Db 1354 GTTGATGGACAGGATATTAGGACCATTAATGTAGGTTTCTACGGGAATCATTTGGTGTG 1413  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyArg 41  
Db 1414 GTGAGTCAGGAACCTGTATTGTTGCCACCATGATAGCTGAAAACATTCGCTATGCCGT 1473  
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaAargGluAlaAsnAlaTyrrPhe 61  
Db 1474 GAAATGTCCACCATGGATGAGATTGAGAAAGCTGTCAAGAAAGCCCAATGCCATGACTTT 1533  
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
Db 1534 ATCATGAACCTGCTCATTAATTTGACACCTGTTGGAGAGAGAGGGGCCCATTTGAGT 1593  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
Db 1594 GGTGGGCAAGCAGAGGATCGCCATTGCGCTGCTGGTTCGCAACCCCAAGATCCTC 1653  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
Db 1654 CTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCAAGCAGTGGTTCAGTGGCT 1713  
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 1714 CTGGATAGCCAGAAAGGTGGACCATTTGTGATAGCTCATCGTTTGTCTACAGTT 1773  
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
Db 1774 CGTAATGCTGACGTCATCGCTGTTTCGATGATGAGTGCATTTGTGGAGAAAGAAATCAT 1833  
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrrTyrrSerLeuValMetSerGln----- 178  
Db 1834 GATGAACCTCATGAAGAGAAAGGCATTTACTTCAAACTTGTCAATGACAGACAGCAGGA 1893  
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrrSerThrGlu 195  
Db 1894 ATGAAGTTGAATTTAGAAATGACGCTGATGAATCCAAAGTGAAATTCATGCTTGGAA 1953  
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209

```
Db 1954 ATGCTCTCAAAATGATTCAAGATCCAGTCTAAATAAGAAAAAGATCAACTCGTAGGAGTGC 2013
Qy 210 Lys--SerAspPheIleAspLysAlaGluSerThrClnSerLysGluIleSerLeu 228
Db 2014 CGTGGATCAACGCCCAAGACAGAAAGCTTAGTACCACAAAGAGGCTGTGAGTAAAGATA 2073
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db 2074 CTTCCAGTTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCTTATTGTT 2133
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2134 GTTGGTGTATTGTGCCATTATAAATGGAGGCTGCAACACGACCAITTTGCAATAATATT 2193
Qy 269 AlaIysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2194 TCAAGATTATAGGGGTTTTTACAAGAAATGTATGATCCTGAAACAAACGACAGAAATAGT 2253
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2254 AACTGTTTTCATATTGTTTTCAGCCCTGGAAATATTTCCTTTTATTACATTTTCCCT 2313
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2314 CAGGGTTTCATATTGGCAAAAGCTGGAGAGATCCTCACCAGCGGCTCCGATACATGTT 2373
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTTPheAspLysGluAsnSerThrGly 347
Db 2374 TTCGGATCCATGCTCAGACAGGATGTGAGTTGTTGTATGATGCCCTAAACACACCTGGA 2433
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2434 GCATTGACTACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGTTCCAG 2493
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2494 CTTGCTGTAATTACCCAGATATAGCAAACTCTTGGGACAGGAATAATATATCTCTCATC 2553
Qy 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2554 TATGGTTGGCAATTAACACTGTTTACTCTTAGCAATTTGTATCCCATCATTTGCAATAGCAGGA 2613
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2614 GTTGTGGAATGAAATGTTGCTGGAACAGCTGAAAGATAGAAAGACTAGAAAGT 2673
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 2674 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTTCTTTGACTCAG 2733
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2734 GAGCAAGATTTGAACATATGTATGTCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2793
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2794 AGGAAGCACACATCTTTGGAATTTACATTTTCTTACCAGCAATGATGATTTTTC 2853
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2854 TATGCTGGATGTTTCGGTTTGGAGCCTACTTGGTGGCACAATAAATCATGAGCTTTGAG 2913
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 2914 GATGTTCTGTAGTATTTCAGCTGTTGCTTGTGGTCCCATGCGCGGCGCAAGTCAGT 2973
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 2974 TCATTTGCTCTGACTATGCAAGGCAAAATATATATATATATATATATATATATATAT 3033
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
```

```
Db 3034 GAAAAAACCCCTTTGATTGACAGCTACAGCAGGAGGCTTAATGCCGAACACATTGGAA 3093
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3094 GGAATGTACATTTGGTGAAGTTGTATTCACTATCCACCCGACCGACATCCAGTG 3153
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3154 CTTAGGAGACTGAGCTGGAGGTGAAGAGGCGCAGACGCTGGCTCTGTGTGGCAGCAGT 3213
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3214 GCGTGTGGGAAGAGACAGCATGCTGTCAGCTCCTCGGAGCGGTTTCCAGCCCTTGGCAGG 3273
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3274 AAAGTGTCTCTGTATGGCAAGAAATAAAGCAGCTGAATGTTCAGTGGCTCCGAGCACAC 3333
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3334 CTGGGCATCGTGTCCAGGAGCCCATCTGTTTGACTGCAGCATTTGCTGAGAACAATTGCC 3393
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3394 TATGGAGACAAACAGCGGGTGTGTCTCACAGGAAGAGATTGTGAGGGCAGCAAGGGAGGCC 3453
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3454 AACATACATGCTTTCATCGAGTCACTGCTCTAATAATAATATAGACTAAAGTAGGAGACAA 3513
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3514 GGAACCTCAGCTCTGTGTGGCCAGAAACCAACGCAATTCCTAGCTCGTGCCTTGTAGA 3573
Qy 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3574 CAGCCTCATATTTTGTCTTGTGTAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 3633
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3634 GTTGTCCAAGAGCCCTGGCAAGCCAGAGAGGCGGCACTGATTTGATTGTTGCTCAC 3693
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3694 CGCTGTCCACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGAATGCGCAGAGTCAAG 3753
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3754 GAGCATGGCAGCATCAGCAGCTGCTGGCAGAAAGGCATCTATTTTCAATGGTCAAGT 3813
Qy 808 AlaGlnSer 810
Db 3814 GTCCAGGCT 3822
RESULT 15
AAZ49332
ID AAZ49332 standard; cDNA; 3860 BP.
XX AAZ49332;
AC AAZ49332;
DT 14-MAR-2000 (first entry)
XX
DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
XX cytokine; wild-type; ds.
OS Homo sapiens.
XX
```

FH Key Location/Qualifiers  
 FT CDS 1..3843  
 FT /\*tag= a  
 FT /product= "Human wild-type MDR-1 protein"  
 FT replace (553..555, GTT)  
 FT mutation  
 FT /\*tag= b  
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given  
 FT in AA249333"  
 XX  
 XX WO961589-A2.  
 XX  
 XX 02-DEC-1999.  
 XX  
 XX 27-MAY-1999; 99WO-US11825.  
 XX  
 XX 28-MAY-1998; 98US-0086988.  
 XX  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 XX Sorrentino B, Bunting K;  
 XX  
 XX WPI; 2000-072615/06.  
 XX P-PSDB; AAY58186.  
 XX  
 XX Ex vivo expansion of hematopoietic stem cells transduced with a  
 XX sequence encoding human multidrug resistance-1, used for bone marrow  
 XX transplantation -  
 XX  
 XX Claim 10; Page 68-70; 113pp; English.  
 XX  
 XX This sequence represents cDNA encoding human wild-type  
 XX multidrug resistance protein MDR-1. MDR-1 is a transmembrane  
 XX efflux pump, responsible for the export of drugs from cells,  
 XX particularly cancer cells. Wild-type MDR-1 shows increased  
 XX resistance to etoposide and decreased resistance to vinca  
 XX alkaloids compared with a mutant form (AAY58187) where the Gly at  
 XX position 185 is replaced by Val. The invention relates to transducing  
 XX hematopoietic stem cells with nucleic acid encoding an MDR protein  
 XX and culturing the modified cells. The modified hematopoietic stem  
 XX cells are useful in bone marrow transplantation (to reconstitute  
 XX hematopoietic systems in patients who have undergone chemotherapy or  
 XX radiation therapy) and in ex vivo gene therapy of genetic defects in  
 XX cells derived from hematopoietic stem cells, e.g., thalassemia,  
 XX Gaucher's disease, sickle cell anemia or leukaemia. The modified  
 XX cells can also be used to identify factors involved in regulating  
 XX proliferation and differentiation in hematopoietic stem cells.  
 XX Hematopoietic stem cells that express MDR-1 will be protected against  
 XX chemotherapeutic agents, so can be engrafted while the patient is  
 XX undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells  
 XX provides sufficient cells to permit standard biochemical analysis.  
 XX Overexpression of MDR-1 allows cytokine-driven expansion of  
 XX hematopoietic stem cells by at least 10-fold compared with a maximum  
 XX of 4-fold in known procedures.  
 XX  
 XX Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 5,558-226 Length: 3860  
 Score: 2436.00 Matches: 466  
 Percent Similarity: 76.79% Conservative: 166  
 Best Local Similarity: 56.62% Mismatches: 177  
 Query Match: 59.72% Indels: 14  
 DB: 21 Gaps: 4  
 US-09-873-409-2 (1-812) x AA249332 (1-3860)  
 QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1357 GTTGATGGACAGGATATTAGGACCATTAATGTAAGGTTCTACGGGAATCATTTGGTGTG 1416  
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1417 GTGAGTCAGGAACCTGTATTGTTGGCCACCAGGATAGCTGAAACACATTCGCTATGGCCGT 1476

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaAa-gGluAAsnAlaTyrAspPhe 61  
 Db 1477 GAAATGTCAACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCCGTATGACTTT 1536  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSer 81  
 Db 1537 ATCATGAACATGCCTCAATAATTTGACACCCCTGGTGGAGAGAGAGGGGCCCATGTGAGT 1596  
 QY 82 GlyGlyGlnLysGlnArgIleAlaAlaAlaLeuValArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1597 GGTGGGAGAGAGCAGAGGATCGCCATTGCGACGTGCGCTGGTTCGCAACCCCAAGATCCTC 1656  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 1657 CTGCTGGATAGGCCACGTCAGCTTGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1716  
 QY 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaAlaHisArgLeuSerThrIle 141  
 Db 1717 CTGGATAAGGCCAGAAAGGTGGACCACTTGTGATGCTCATCGTTTGTCTACAGTT 1776  
 QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
 Db 1777 CGTAATGCTGACGTCATCGCTGGTTCGATGATGAGGATCATTTGTGAGAAAGAAATCAT 1836  
 QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
 Db 1837 GATGAACCTCATGAAAGAGAAAGGCATTTACTTCAAACTGTTCACAAATGCAGACAGCAGGA 1896  
 QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
 Db 1897 AATGAAGTTGAATTAGAAATGCAGCTGATGAATCAAAAGTGAAATGATGCCCTGGAA 1956  
 QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
 Db 1957 ATGCTTCAATGATTCAAGATCCAGTCTAATPAGAAAAGATCAACTCGTAGGAGTGT 2016  
 QY 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluSerLeu 228  
 Db 2017 CGTGATCACAAGCCCAAGACAGACAGAAAGCTTAGTACAAAGAGGCTCTGGATGAAATATA 2076  
 QY 229 ProGluValSerLeuLysLysIleLysLysLeuAsnLysProGluTyrProPheValVal 248  
 Db 2077 CCTCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTTATTTGTT 2136  
 QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
 Db 2137 GTTGGTGTATTTTGGCCATTATAATGGAGCCCTGCACACAGCATTTGCAATATATTT 2196  
 QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287  
 Db 2197 TCAAGATTATAGGGGTTTTCACAGAATTGATGATCCTGAAACAAACAGCAGATACT 2256  
 QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
 Db 2257 AACTTGTGTTTCACTATTGTTCTAGCCCTTGGAAATATTCTTTTATTACATTTTTCTCT 2316  
 QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
 Db 2317 CAAGGTTTCACATTTGGCAAGCTGGAGATCCTCACCAGCGGCTCGGATACATGTT 2376  
 QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
 Db 2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCTTAAACACACCACTGGA 2436  
 QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
 Db 2437 GCATTGACTACCAAGCTGCCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAG 2496  
 QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
 Db 2497 CTTGCTGTAATTAACCAGAAATATACCAAACTCTGGACAGGAAATATATATATCCTTCATC 2556



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 5585.08 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-2  
Perfect score: 4079  
Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115420\_19240/app\_query.fasta\_1.7544  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CCN 1 1 30544 @runat\_27032003\_115420\_19240 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcti:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcti:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1008	24.7	943	9	AL520322	AL520322 AL520322
2	834	20.4	998	14	BM904842	BM904842 AGENCOURT
3	780	19.1	2676	11	AK014319	AK014319 Mus muscu
4	770.5	18.9	1019	12	BG248052	BG248052 602359987
5	753	18.5	726	12	BG293345	BG293345 602390738
6	751	18.4	944	12	BF796582	BF796582 602358463
7	743.5	18.2	886	17	BH139685	BH139685 ENTNA477R
8	740.5	18.2	871	17	AZ682350	AZ682350 ENTKB16TF
9	734.5	18.0	932	17	AZ670821	AZ670821 ENTJN69TF
10	728.5	17.9	947	17	AZ683753	AZ683753 ENTIL96TF
11	726.5	17.8	886	17	AZ540627	AZ540627 ENTQ18TF
12	723.5	17.7	880	17	AZ687805	AZ687805 ENTJLJ52TF
13	721.5	17.7	939	14	BQ720763	BQ720763 AGENCOURT
14	720.5	17.7	897	17	AZ541090	AZ541090 ENTDS67TR
15	711.5	17.4	913	17	BH155700	BH155700 ENTRO54TR
16	709.5	17.4	834	17	AZ548312	AZ548312 ENTG07TR
17	708.5	17.4	853	17	AZ679807	AZ679807 ENTH167TR
18	705.5	17.3	823	17	AZ532602	AZ532602 ENTCSR50TR
19	697.5	17.1	861	14	BQ717101	BQ717101 AGENCOURT
20	691	16.9	780	12	BG587938	BG587938 EST489713
21	690.5	16.9	899	17	BH154857	BH154857 ENTRO47TF
22	689.5	16.9	891	17	AZ682250	AZ682250 ENTK168TR
23	688	16.9	1123	11	AY108485	AY108485 Zea mays
24	683.5	16.8	1341	11	AY108285	AY108285 Zea mays
25	683	16.7	600	13	BM486593	BM486593 pgm2n.pk0
26	682.5	16.7	919	17	AZ690701	AZ690701 ENTIV68TR
27	678	16.6	926	14	BQ123477	BQ123477 EST609053
28	677.5	16.6	795	12	BF313560	BF313560 601900192
29	669	16.4	547	13	BQ303071	BQ303071 BJ030711
30	667	16.4	937	17	BH150760	BH150760 ENTQ042TR
31	663	16.3	835	13	BJ349604	BJ349604 BJ349604
32	662	16.2	636	10	AV962688	AV962688 AV962688
33	661	16.2	747	13	BJ445774	BJ445774 BJ445774
34	659.5	16.2	628	12	BG080311	BG080311 H3052B06-
35	659	16.2	852	17	BH720383	BH720383 BOHZ294TF
36	657	16.1	715	14	BQ869512	BQ869512 QGD6G09.Y
37	657	16.1	785	10	AV709991	AV709991 AV709991
38	655.5	16.1	787	12	BG584063	BG584063 EST485823
39	655	16.1	919	14	BQ724901	BQ724901 AGENCOURT
40	653	16.0	609	10	AV986144	AV986144 AV986144
41	651.5	16.0	921	17	AZ687628	AZ687628 ENTIR96TR
42	651	16.0	746	13	BJ380617	BJ380617 BJ380617
43	650	15.9	729	13	BJ353293	BJ353293 BJ353293
44	647	15.9	1813	11	U66688	U66688 Homo sapien
45	641	15.7	820	13	BJ356540	BJ356540 BJ356540

# ALIGNMENTS

RESULT 1  
AL520322  
LOCUS  
DEFINITION AL520322 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5  
prime, mRNA sequence.  
ACCESSION AL520322  
VERSION AL520322.1 GI:12783815  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 943)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization

[illegible]

```

EST.
human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 998)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@email.nih.gov
              Tissue Procurement: ATCC/DCTP
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12279 row: f column: 24
              High quality sequence stop: 738.
FEATURES
    source
    1..998
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5557655"
        /clone_lib="NIH MGC 72"
        /tissue_type="melanotic melanoma"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: Not
              Site 2: Sali; Cloned unidirectionally. Primer: Oligo
              Average insert size 2 kb. Library constructed by Life
              Technologies."
    BASE COUNT      271 a      223 c      248 g      256 t
    ORIGIN
Alignment Scores:      2,2e-86      Length:      998
Pred. No.:             834.00      Matches:      164
Score:

```



Percent Similarity:	74.12%	Conservative:	68
Best Local Similarity:	52.40%	Mismatches:	77
Query Match:	20.45%	Indels:	4
DB:	14	Gaps:	0
US-09-873-409-2 (1-812) x BM904842 (1-998)			
QY	335	AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla	354
DB	61	GATGTGAGTTGGTTGATGACCCCTAAAAACACCACTGGAGCAATGACTACCGGCTGCC	120
QY	355	IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn	374
DB	121	AATGATGCTCTCAAGTTAAAGGGCTTATAGTTCAGGCTTGCTGTAATTAACCAAGAT	180
QY	375	AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe	394
DB	181	ATAGCAATCTTGGACAGGAGTAATATATCTTCACTATGTTGGCAACTAACACTG	240
QY	395	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	414
DB	241	TTACTCTTAGCAATTTGACCATCATTTGCAATAGCAGGAGTTGTGAATCAAAATGTTG	300
QY	415	ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu	434
DB	301	CTGGCAAGCACTGAAAGATAGAAAGAACTAGAAGGCTGCTGGGAAGATCGCTACTGAA	360
QY	435	AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet	454
DB	361	GCAATAGAAACCTCCGACCGTGTCTTCTTGACTCAGGAGCAGAGTTTGAACATATG	420
QY	455	TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly	474
DB	421	TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGCACACATCTTGGGA	480
QY	475	SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe	494
DB	481	ATTACATTTTCTTCCACCGCAATGATGATTTTCTTATGCTGATGATGTTTCCGGTTT	540
QY	495	GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr	514
DB	541	GGAGCTTCTTGGCGGACATAAATCATGAGCTTGGAGATGTTCTGTAGTATTTTCA	600
QY	515	AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer	534
DB	601	GCTGTTCTTCTTGGTGCATGGCGTGGGCAAGTCACTTCTTCTGCTGCTGATGCTCC	660
QY	535	LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp	554
DB	661	AAAGCCAAATATATCAGCAGCCCATCATCATGATCATATGTAATGAAATCCCTTTGATTGAC	720
QY	555	SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu	574
DB	721	AGTACAGCAGGAGGAGGCTTAATGCCGACACATTTGAAGAAATGTCACATTTGGTGA	780
QY	575	ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe	594
DB	781	GTGTATTCAACTATCCACCGACCGACATCCAGTCTTCAGGACTGAGGCCCTTGGGA	840
QY	594	rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS	614
DB	841	GGTGAAGAGGGGCGAGCCCTTGGTCTGTTGGGGCAGCAGTGGCTGTGGGAAGACACAG	900
QY	614	er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlyValLeuPheAspGly	633
DB	901	TGGGTCCAGCTCTGGAGCGGGTCTAACACCCCTTGGCAAGAAATGCTGCTTGGTGGC	960
QY	634	ValAspAlaLys-GluLeuAsnValGlnTyr 643	
DB	961	AAAAGATNAACCACTGGATGTTCCATGG 991	
RESULT	3		
AK014319			

LOCUS	AK014319	2676 bp	mRNA	linear	HTC 19-JAN-2002
DEFINITION	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:322401P09:ATP-binding cassette, sub-family B (MDR/TAP), member 8, full insert sequence.				
ACCESSION	AK014319				
VERSION	AK014319.1	GI:12852089			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library clone:322401P09.				
ORGANISM	Mus musculus				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	2				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)				
PUBMED	20530913				
REFERENCE	4				
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamilya, M., Lee, N.H., Lyons, P., Mombert, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo, O., K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
MEDLINE	Nature 409 (6821), 685-690 (2001)				
PUBMED	21085660				
REFERENCE	5 (bases 1 to 2676)				
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Iwatsuki, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,				



Db 1200 TTCAATGGGGCTCCCTGTGGCTGGACAGCAGCTGAAGGGGAGACCTCATGTCTCTC 1259  
Qy 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyIleThrLeuValLeuAla 530  
Db 1260 CTGGTGGCTCCAGACAGTACAGAGTCTATGGCC-----AGCTCTCTGCTCTGT 1313  
Qy 531 ProGluTyrSerIlysalYsSerGlyAlaAlaHleLeuPheAlaLeuGluIlyLys 550  
Db 1314 GGTCAAGTGTACGTGGCTGAGTGGGGAGCCGAGTCTTCGAATACATACAGCTGAGC 1373  
Qy 551 ProAsnIleAspSerArgSerGlnGluGlyIlySerProAspThrCysGluGlyAsnLeu 570  
Db 1374 CCGTGTATCCCATTCACCGGGGGCTACTGTCATCCCAACAAAGGACATTCGTGGTTCATC 1433  
Qy 571 GluPheArgGluValSerPheThrProCysArgProAspValPheIleLeuAsnGly 590  
Db 1434 ACCTTCMAAATGTCACTTCACCTACCTCCCTGCAGACCTGGCTCAATGGCTCAAGGAC 1493  
Qy 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610  
Db 1494 TTCACCTGAAGCTGCCTCTGCAAGATTGTGGCTCTTGTGGCCAGTCTGGGGAGGA 1553  
Qy 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeu 630  
Db 1554 AAGACCACAGTTCGCTCTGTCTGCAACGCTTCTATGACCTCAAGCTGGCTCGGTGAGC 1613  
Qy 631 PheAspGlyValAspAlaIlyGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 649  
Db 1614 TTGATGGGATGACCTGCGCACTCTCAACCTCTCGGCTCGGGCCAGGTCTAGGT 1673  
Qy 650 IleValProGlnProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669  
Db 1674 TTCAATCAGCAGCAGCAGTCTCTGTTTGAACAACCATCATGAGAATAATCCGATTTGGG 1733  
Qy 670 -----AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687  
Db 1734 AAGCTGATCTTCCGAT-----GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781  
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707  
Db 1782 AATGCCACAGTTCATCAGCAGCTTCCCGATGGCTACAGCACTGTGGTTGTGACCGG 1841  
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727  
Db 1842 GGCACAACCTTGTCTGTGGCCAGACAGCGCTACCCATCGCACGTGCCCTCATCAAG 1901  
Qy 728 LysProIlyIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747  
Db 1902 CAGCCACACAGTGTATCTCTGGACAGGCCACCAAGTCCGTAGATGCAGAATCCGAGAGG 1961  
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
Db 1962 GTGGTACAGAGGCCCTTGGACCGCGCAGTGTCTGGCCGACCGGTGTGGTTCATGGCCAC 2021  
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAenGlyLysIleLys 787  
Db 2022 CGCTTAGTACTGTCCGTGAGCCCACTCCATCATTTGTATGGCCATGGCCAACTGTGT 2081  
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807  
Db 2082 GAGGCTGGGACCCACAGAAGAACTCTTAAAAAAGCGGGCTCTATTTCAGAGCTTATCCGG 2141  
Qy 808 AlaGlnSerVal 811  
Db 2142 AGACAAACCCCTG 2153  
RESULT 4  
BG248052 1019 bp mRNA linear EST 13-FEB-2001  
LOCUS 60235987F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',  
DEFINITION mRNA sequence.  
ACCESSION BG248052

VERSION BG248052.1 GI:12757867  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 1019)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10134 row: n column: 21  
High quality sequence stop: 650.  
FEATURES  
Location/Qualifiers  
1..1019  
Source  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4488404"  
/clone\_lib="NCI CGAP Mam1"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="10 months, virgin"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"  
BASE COUNT 276 a 262 c 283 g 196 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.2e-79 Length: 1019  
Score: 770.50 Matches: 154  
Percent Similarity: 79.17% Conservative: 36  
Best Local Similarity: 64.17% Mismatches: 47  
Query Match: 18.89% Indels: 3  
DB: 12 Gaps: 1  
US-09-873-409-2 (1-812) x BG248052 (1-1019)  
Qy 567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586  
Db 4 GAAGGAAATGTGCAATTTAGTGGAGTCTGTTCAACTATCCACCCGACCCAGCATCCCA 63  
Qy 587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyIleThrValAlaPheValGlySer 606  
Db 64 GTCTCTCAGGGGTGAGCCTTGAGGTGAAGAGGGCCAGACGCTGGCCCTGGTGGGGCAGC 123  
Qy 607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626  
Db 124 AGTGGCTGGGGAGAGACACAGTGGTCCAGCTCTCGAGGCGCTTCTACGACCCCATGGCT 183  
Qy 627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646  
Db 184 GGATCAGTGTCTTAGATGGCAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243  
Qy 647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666  
Db 244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTTCTTGTACTGCAGCATTCGAGAGAAACATT 303  
Qy 667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 686  
Db 304 GCCTACGGAGACAAACAGCCGGTCTGTCTTATGAGGAGATTGTGAGGGCAGCCAAAGAG 363  
Qy 687 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706

```

Db 364 GCCAACATCCACCACTTCATCGACTCGTACTGATAAATACACACAGTAGGAGAC 423
QY 707 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArgAlaLeuLeu 726
Db 424 AAGGGCACTCAGCTGTCTGGGTGGGCAAGACGCGCATGCCATCGACGCGCTCGTC 483
QY 727 GlnLysProLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 746
Db 484 AGACAGCTCACATTTCCTCTGGACAGCAACATCAGCTCTGGATACAGAAAGTAA 543
QY 747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766
Db 544 AAGTTGTCCAGGAAGCGCTGGACAAAGCCAGGAGCGACCTGCATTGTGATCGCT 603
QY 767 HisArgLeuSerAlaLeuGlnAsnAlaAspLeuLeuLeuValValLeuHis-AsnGlyLysI 786
Db 604 CACCGCTGTCCACCATCCCAAGACCGCGACTTGTATCGGTGGTGATTCAAGAACCGGAAG 663
QY 786 LeLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyPhe 803
Db 664 TCAAGGAGCAGCGGCCACCCCAACAGCAAGTGGTGGCGCAGAGGGCATCTACTTC 719

RESULT 5
BG293345
LOCUS 602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
DEFINITION mRNA sequence.
ACCESSION BG293345
VERSION BG293345.1 GI:13052943
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 726)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10371 row: 1 column: 09
High quality sequence stop: 658.

FEATURES
source
1..726
/organism="Mus musculus"
/db xref="taxon:10090"
/clone="IMAGE:4502552"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 189 a 220 c 190 g 127 t
ORIGIN

Alignment Scores:
Pred. No.: 3,9e-77 Length: 726
Score: 753.00 Matches: 152
Percent Similarity: 84.29% Conservative: 25
Best Local Similarity: 72.38% Mismatches: 30
Query Match: 18.46% Indels: 3
DB: 12 Gaps: 0

```

```

US-09-873-409-2 (1-812) x BG293345 (1-726)
QY 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
Db 2 CTGGCCCTGTGGGAGCAGTGGCTGGGGAAGAGACAGTGTCTCAGCTGCTCGAGCGC 61
QY 621 LeuTyAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
Db 62 TTTATGACCCCATGCTGGATCAGTCTCTTAGATGGTCAAGAGCAAGAACAATCAAT 121
QY 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
Db 122 GTCAGTGGCTCGAGTCAACTGGGCATTGTCTCCAGGAACCCATTCTCTTTGACTGC 181
QY 661 SerIleAlaGlnLeuAlaTyGlyAspAsnSerArgValValProLeuAspGluLeu 680
Db 182 AGCATCGAGAGAACATCGCCTATGGAGAACACAGCGGGTCTGTCCTCATGTAGATT 241
QY 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTy 700
Db 242 GTGAGGCGACCAAGGAGGCCAACATCCACCTTTCATCGAGACGCTGCCCAAAATAT 301
QY 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 302 AACACAAGAGTAGGAGACAAAGGGGAGCGAGCTCTCTGGGGGCCAGAGCAGAGATTGCC 361
QY 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 362 ATCGCCCGAGCCCTCATCAGACAGCTCGGGTCTCTACTGCTGGATGAAGCCAGTCACT 421
QY 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 422 CTGGATCTCAGAGTGAAGAAGTGTTCAGGAAGCAGCTGGACAAAGCCAGGAGCGCG 481
QY 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 780
Db 482 ACCTGCATTCCGATCGTCTACCGCTGTCCACCATCCAGAACGCGGCACCTTGATCGTGT 541
QY 780 LLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 800
Db 542 GATTGAGACGGCAAGGTCAAGGAGCAGCGCACCCACAGCAGCTGCTGGCGCAGAAGG 601
QY 800 spIleTyPhe-LysLeuValAsn 807
Db 602 GCATCTATTCTCAACTGGTCAAC 625

```

```

RESULT 6
BF796582
LOCUS 602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
DEFINITION mRNA sequence.
ACCESSION BF796582
VERSION BF796582.1 GI:12101636
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 944)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D.; Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9955 row: n column: 15

```



```
Db 20 GCAAAGTACTTATGATGTTATTGATAGAAACCACT-AITGATGTTATAGTGAAGAA 78
Qy 560 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 579
Db 79 GGTGAACATTAAATGATGTTAAAGGTGAATTAAGATTAAGATTTGTTTCAGATAT 138
Qy 580 ProCysArgProAspValPheLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 599
Db 139 CCAACAGACCAACAATCTGCTTGAAGGTATTTCAATCAAGTAGAACAAGAA 198
Qy 600 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 619
Db 199 ACTGTTGATAGTAGGAGCATCAGATGTGTAATCAACATCAGTTCAGTTGATTGAA 258
Qy 620 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 639
Db 259 AGATTTTATGATCCAACACATCGGAGATGTTATTTAGATGACATAATATCAAAAGATTG 318
Qy 640 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 659
Db 319 AATATTTCATTTCTTAAGAGTCAAAATGGAATGGTAGGACCAAGACCAAGTATTATTGCT 378
Qy 660 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 678
Db 379 GAAGTGTATGGATAATATTAGAGAGGAGTACTTAAGAGGTTGAAGTAAGTAATGAA 438
Qy 679 GluLeuLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGlu 698
Db 439 CAAATTTATGCTGCTGTCTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 498
Qy 699 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 718
Db 499 GGATATAACAATGTTAGGTAGTAGAGTGCACAAATTTTCAGGAGGACAAACAAAGA 558
Qy 719 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 738
Db 559 ATTGCTATTGTCACGTCATGATTAGAAATCCAAAGGTTATTACTCGATGAAGCTACA 618
Qy 739 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 758
Db 619 TCAGCATTGATTTCAGAAAGTGAAGATTTGTACAAAGATGCATTCGACAAAGACGAAA 678
Qy 759 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 778
Db 679 GGAAGAACAAATTTGATTCACATAGATATCACTATTCAAAATGCAGATCAATA 738
Qy 779 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 798
Db 739 TGTGTTATTATCAGAGGAAGATTCAGAAAGAGGACACATCAAGAGTATTAGATTG 798
Qy 799 ArgAspIleTyrPheLysLeuValAsnAlaGln 809
Db 799 AAAGGATTTATTATACACTTGTCTATGCAACAA 831

RESULT 8
AZ682350 871 bp DNA linear GSS 14-DEC-2000
LOCUS ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION AZ682350
VERSION AZ682350.1 GI:11819496
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 871)
AUTHORS Eukaryota; Entamoebidae; Entamoeba.
TITLE Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
JOURNAL HMI:IMSS sheared DNA library
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
```

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjlottus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library

Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 16  
High quality sequence stop: 860.

FEATURES  
source  
1. .871  
/organism="Entamoeba histolytica"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/notes="Vector: pHS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, D.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

BASE COUNT 349 a 117 c 174 g 231 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.55e-75 Length: 871  
Score: 740.50 Matches: 147  
Percent Similarity: 70.04% Conservative: 47  
Best Local Similarity: 53.07% Mismatches: 82  
Query Match: 18.15% Indels: 1  
DB: 17 Gaps: 1

US-09-873-409-2 (1-812) x AZ682350 (1-871)

Qy 534 SerLysAlaLysSerGlyAlaAlaHisPheAlaLeuLeuGluLysLysProAsnIle 553  
Db 39 AATAGTCTAAAGTTCGCGCATTTAATGTTTATCAACAATTTGATAGATCCAGATATT 98  
Qy 554 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 573  
Db 99 GATTGTCAGTCTATTGGAGGTGAATGTCCAACCTCAGTGTAAATGGAATATTAGATTGAG 158  
Qy 574 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 593  
Db 159 GATGTTCAATTTGTTTATCCAAACAGAGCTGTCTCATCATGTATTATAAAGGACTTGACCT 218  
Qy 594 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 613  
Db 219 GAAATTAAAGAAAGGACAAACAATTCATTTAGTTCAGATCAGGATGTTGGAAGTCAACT 278  
Qy 614 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 633  
Db 279 ACTATTCAATTAATCCAAAGAAATTAATGATCCAAATGGTGAAGAGTAACATTAGACGA 338  
Qy 634 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 653  
Db 339 AAGATATACGAGAGTTGAATATCAATGTTTAAAGAAATCAATAGGATTAGTTGGACAA 398  
Qy 654 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 673  
Db 399 GAACAGTGTGTTTTCAGGAGAACAAATTCGAGAAATATTATGCTTGGAGCTAAAGAGGA 458  
Qy 674 ValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 692



Db 459 GCACACCAAGTGAAGAGGATGATTGAATGTGCTAAATGCAATGCACATGACTTC 518  
Qy 693 ileGlulGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 712  
Db 519 ATTCTCTAACTCCAGAGGATATGACACAATAATTTGGAGAAAAGGAGCATTTATATCA 578  
Qy 713 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 732  
Db 579 GGAGGACAAACAAAGAAATGCAATTTGACGTCATTTGATTCGAAACCCATCTATTCTT 638  
Qy 733 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 752  
Db 639 CTCTTGATGAAGCTACATCAGCAGCTTGATACACAAGTGAAAGATTTGACAAAGACA 698  
Qy 753 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 772  
Db 699 CTTGAAAAGAGCTCTAAAGGAAGACAACAATTTATTTGTAGCACATAGACTTAACAAC 758  
Qy 773 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHis 792  
Db 759 AGAAATGCGATATAATATGTGTATTTCATCAAGGAGAAATTTTGAACAAAGGAAACAT 818  
Qy 793 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809  
Db 819 CAAGAATTAATGATTGTAAGGAACAATATTATGGATTAGTCAAAAGACAA 869

RESULT 9  
LOCUS AZ670821/c 932 bp DNA linear GSS 14-DEC-2000  
DEFINITION ENTJUN69TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.  
ACCESSION AZ670821  
VERSION AZ670821.1 GI:11807967  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica.  
ORGANISM Entamoeba histolytica.  
REFERENCE 1 (bases 1 to 932)  
AUTHORS Eukaryota; Entamoebidae; Entamoeba.  
TITLES Loftus, B., Van Aken, S. and Fraser, C.  
Determination of clone end sequences from Entamoeba histolytica  
HM1.IMSS sheared DNA library  
Unpublished (2000)  
CONTACT: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjoftus@tigr.org  
DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 46  
High quality sequence stop: 867.  
Location/Qualifiers  
1. .932  
/organism="Entamoeba histolytica"  
/strain="HM1:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

BASE COUNT 258 a 182 c 125 g 367 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,76e-75 Length: 932  
Score: 734.50 Matches: 150  
Percent Similarity: 71.11% Conservative: 42  
Best Local Similarity: 55.56% Mismatches: 77  
Query Match: 18.01% Indels: 2  
DB: 17 Gaps: 1  
US-09-873-409-2 (1-812) x AZ670821 (1-932)  
Qy 533 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsn 552  
Db 835 TTTTGCTACAGCTTAAGCTTCTGCATATAGAAATTTATCAAAACAATTCATAGATCCAGAT 776  
Qy 553 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 572  
Db 775 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAAATGGAATATTACATTA 716  
Qy 573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592  
Db 715 GAAGATGTTCAATTTAGATATCCACAAGACCACCTAAACAATAATCTTGGTGGACTTGT 656  
Qy 593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612  
Db 655 CTTGAAATTTAAGAAAGGACAAACAGTTGCTATTAGTAGGAGCATCAGATGTGTAATATCA 596  
Qy 613 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632  
Db 595 ACTACTATTCAATTTAGTCCAAAGAAATTTATGATCCCAAGTTGGTGGGTCAATATAGAT 536  
Qy 633 GlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValPro 652  
Db 535 GGAAAGATTTAAGAGATTTAAATATCAATGTTAAGAAATCAATAGATGATTGTGGA 476  
Qy 653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 672  
Db 475 CAAGAACCTATTTTGTTGCTATGCTACTATTAGAGAAATATTATGCTGGAGCTAGAGAT 416  
Qy 673 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaIleHisSer 691  
Db 415 GGAGAAACACCACTGAAGAGAGATGATTGAATGTCTAAATATGCAATGCACATGAA 356  
Qy 692 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 711  
Db 355 TTTATTCTTCATCTTCCAGAGGATATGATACAATGTTAGGAGAAAAGAGCTGCATTA 296  
Qy 712 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 731  
Db 295 TCAGGAGGACAAACAAAGAAATGCTATTGACGTCGCTATTGATTAGAAACCTACAAT 236  
Qy 732 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAspSerGluLysValValGlnHis 751  
Db 235 TTATTACTTGTATGAGCTACATCAGCAGCTTGATACACAAGTGAAGATTTGTACACAA 176  
Qy 752 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 771  
Db 175 GCATTTGAAAAGAGCTAGTCAAGGAGAACACAAATTTGTTGTAGCACATAGATTAACT 116  
Qy 772 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThr 791  
Db 115 GTTAGAAATGCAAGTAGAATTTCTGTATTCCATCAAGGAGAAATTTAT-GAACAAGGACA 57  
Qy 792 HisGlnGluLeuLeuArgAsnArgAspIle 801  
Db 56 CATCAAGATTAAGGATTTAAAGAGGATTTA 27  
RESULT 10  
LOCUS AZ683753 947 bp DNA linear GSS 14-DEC-2000

DEFINITION ENT196TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.

ACCESSION AZ683753

VERSION AZ683753.1 GI:11820899

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 947)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 40

High quality sequence stop: 778.

Location/Qualifiers

1..947

/organism="Entamoeba histolytica"

/strain="HM1:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/notes="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD

Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 379 a 121 c 178 g 269 t

ORIGIN

Alignment Scores:

Pred. No.: 4.54e-74 Length: 947

Score: 728.50 Matches: 149

Percent Similarity: 69.82% Conservative: 43

Best Local Similarity: 54.18% Mismatches: 81

Query Match: 17.86% Indels: 2

DB: 17 Gaps: 1

US-09-873-409-2 (1-812) x AZ683753 (1-947)

Qy 536 AlaLysSerGlnAlaHisLeuPheAlaLeuGlnLysLysProAsnLeuSer 555

Db 69 GCAAAAGCATCAGCCTATAAATATTTACACAATTCATAGATCCAGATATTGTT 128

Qy 556 ArgSerGlnGluLysLysProAspThrCysGluGlnAlaLeuGluPheArgGluVal 575

Db 129 CAGTCTATTGGAGTGAAATGTCACACTGAGTGAATGGAATATTAGATTGATGTT 188

Qy 576 SerPheThrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerile 595

Db 189 CAATTGTTTATCCACAGACCATCTCATCATGTATTAAAGGAATTCACITGAAATT 248

Qy 596 GluArgGlyLysThrValAlaPheValGlySerGlyCysGlyLysSerThrSerVal 615

Db 249 AAGAAAGGAGAATAAATGCTATTAGTTGGAGCATCAGGATGTAGAAAGTCAACTACTATT 308

Qy 616 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlnValValLeuPheAspGlyValAsp 635

Db 309 CAATTAATCAAAGAAATATGAAACCAATATGGTGGAGAGTAGTACATTAGATGGAAAGAC 368

Qy 636 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 655

Db 369 ATAAGAGAGCTTGAATATCAATGGTTTAAAGAAATCAATAGGAATAGTTGGACAAAGACCA 428

Qy 656 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly---AspAsnSerArgVal 674

Db 429 GTATTATTTTCAGGAACAATTCGAGAAATATCATCTCTGGAGCTTAAGAGAGAGAAACA 488

Qy 675 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 694

Db 489 CTAAGTAAAGAGAGATGATTGAATCTACTAAATGGCAATGCAATGCATGACTTATTCTTCT 548

Qy 695 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 714

Db 549 AAACCTCCAGAGAGATATGACACAAATAATTTGGAGAAAAAGGAGCATTTATTATCAGAGGA 608

Qy 715 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 734

Db 609 CAAAAACAAGAAATTCGAATTCGACGTGCTGATTGATTGCGAAACCATCTATTCTTCTCT 668

Qy 735 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 754

Db 669 GATGAAGCTACATCAGCACTTGATACACAAGTGAAAGATTTGTACAAGAAGCACTTGAC 728

Qy 755 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 774

Db 729 AAAGCATCAAAAGGAAGAAACAACAATTTATTGTAGGACATAGACTATTCGACTATTTCAA 788

Qy 775 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 794

Db 789 GCAGATCAATATGTTTATTATGAGAGGAAAAAATAGTAGACAAGAAGACACATCAAGAA 848

Qy 795 -LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 808

Db 849 TTTATTGGATTTGAAAAGATTTTATTATACATCTGCTTATGCA 891

RESULT 11

LOCUS AZ540627

DEFINITION ENTQ18TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.

ACCESSION AZ540627

VERSION AZ540627.1 GI:11147603

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 886)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica HM1:IMSS sheared DNA library

JOURNAL Unpublished (2000)

COMMENT Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 23

High quality sequence stop: 856.

Location/Qualifiers

1..886

/organism="Entamoeba histolytica"

FEATURES

source

RESULT 12					
AZ687805/c					
LOCUS					
DEFINITION	AZ687805	880 bp	DNA	linear	GSS 14-DEC-
	ENTLU52Ff	Entamoeba histolytica	Sheared	DNA	Entamoeba histolytica
	genomic.	DNA sequence.			

GSS.  
 07605.1 GA.11024951  
 KEYWORDS  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica  
 Eukaryota; Entamoebidae; Entamoeba.  
 1 (bases 1 to 880)  
 Loftus, B., Van Aken, S. and Fraser, C.  
 Determination of clone end sequences  
 from Entamoeba histolytica

**JOURNAL**  
UNPUBLISHED (2000)

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 6200  
 Fax: 301 838 3543  
 Email: [bjloftus@tigr.org](mailto:bjloftus@tigr.org)  
 Clones are derived from the *Entamoeba histolytica* HM1:IMSS strain

Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 30

```

FEATURES
  source
    Location/Qualifiers
      1. .880
      /organism="Entamoeba histolytica"
      high quality sequence exp. exp.

```

```

/clone_lib="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pUC19. Site 1: BamHI. Constructed at
/submitter="EMBL:INMSS"

```

Genomic DNA isolated from broth cultures of *E. histolytica* using a method described by Clark and Diamond (Clark and Diamond 1982). Genomic histolysates

method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used

H.O. and Venter, J.C. (Making small insert libraries whole genome shotgun sequencing projects. In *Genome Sequencing: A Practical Approach*, eds. M. Vaudin and

BASE COUNT	247 a	167 c	122 g	344 t
ORIGIN				

Pred. No.:	1.54e-73	Length:	880
Score:	723.50	Matches:	151
Percent Similarity:	69.93%	Conservative:	42

Query Match:	17.74%	Indels:	2
DB:	17	Gaps:	1



Db 423 ATGTATGTCAGAGTTTGCAGGTACCATACAGAACTCTTTTGAGGAAAGCACACATCTTT 482  
Qy 474 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 493  
Db 483 GGAATTACATTTCTTCCACCGGCAATGATGATATTTTCTATGCTGGATGTTCCGG 542  
Qy 494 PheGlyAlaTyrIleuLeuGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 513  
Db 543 TTTGGAGCCTACTTGTGGGCACATAAACTCATGAGCTTTCAGGATGTTCTGTTAGTATT 602  
Qy 514 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrIleuValLeuAlaProGluTyr 533  
Db 603 TCAGCTGTGTTCTTGGTCATGGCCGTCGAGGCAATGAGTTCATTTGCTCTCGACTAT 662  
Qy 534 Ser-LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn11 553  
Db 663 GCCCAAGCCAAATATACAGCCACATCATCATGATCATTTGAAAAACCCCTTTGAT 722  
Qy 553 eApeSerArgSer-GlnGluGlyLysLysPheProAspThrCysGluGlyAsnLeuGlu---P 572  
Db 723 TGACAGCTACAGCACCGGAGGCTAATGCCGAACACATTTGGAAGGAATGTCCACATTT 782  
Qy 572 heArgGluValSerPhePheTyrPro 580  
Db 783 GGTGAAAGTTGATTTCAACTATATCCC 808  
RESULT 14  
AZ541090 897 bp DNA linear GSS 14-NOV-2000  
LOCUS ENTDS67R Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION genomic, DNA sequence.  
ACCESSION AZ541090  
VERSION AZ541090.1 GI:11148493  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica.  
ORGANISM Entamoeba histolytica  
REFERENCE 1 (bases 1 to 897)  
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
JOURNAL HM1:IMSS sheared DNA library  
COMMENT Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjl@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 81  
High quality sequence stop: 890.  
Location/Qualifiers  
1. 897  
/organism="Entamoeba histolytica"  
/strain="HM1:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHOSt1; Site 1; Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.

BASE COUNT 335 a 110 c 177 g 275 t  
ORIGIN  
Alignment Scores: 3.58e-73 Length: 897  
Score: 720.50 Matches: 143  
Percent Similarity: 74.59% Conservative: 39  
Best Local Similarity: 58.61% Mismatches: 61  
Query Match: 17.66% Indels: 1  
DB: 17 Gaps: 1  
US-09-873-409-2 (1-812) x AZ541090 (1-897)  
Qy 567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586  
Db 94 AAAGGTGAATTTGAATTTAAAGACATTTGTTTCAGATATCCAAAGACACAGCAATTCT 153  
Qy 587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606  
Db 154 GTCTTTGAAAGGTATTTCATTTCAAAGTAGAACCAAGGAAAAAAGTCTTGCATTAGTAGGAGCA 213  
Qy 607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626  
Db 214 TCAGGATGTGGTAAATCAACATCAGTTGATTTGAAAGATTTTATGATCCACACAT 273  
Qy 627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646  
Db 274 GAGATGATATTATAGATGACATATATATCAAGATTTTGAATATTCATTTCTTAAGAAGT 333  
Qy 647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666  
Db 334 CAAATTGGAATGTAGGACCAAGAACCCAGTATTATTGCTGAAAGTGTATTGGATAATATT 393  
Qy 667 AlaTyrGlyAspAsnSerArgVal---ValProLeuAspGluIleLysGluAlaAsn 685  
Db 394 AGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAACAAATTTATGCTGCTGCTAAA 453  
Qy 686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705  
Db 454 ATGGCAATGCAATGACTTTATTTTCAGCAATGCCAGAGGATATTAACAATGTTAGGT 513  
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725  
Db 514 GATAGAGGTGCACAAATTTTCAGGAGGACAAAAACAAAGAAATGCTATTGACGCTGATTG 573  
Qy 726 LeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745  
Db 574 ATTAGAAATCCAAAGATGTTATTACTCGATGAAGCTACATCAGCACTTGATTCAGAAAGT 633  
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765  
Db 634 GAAAGATTGTACAGATGCACTTTGACAAAGCAGCNAAGGAGGAAGAACAAATTTGAATT 693  
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlyLys 785  
Db 694 GCACATAGATTATCAACTATTCAAATATGCAGATCAATATGTTGTTATTATGAGAGGAAGA 753  
Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805  
Db 754 ATTGCAAGAGGAGGACACATCAAGAGTTATTAGATTTTGAAGGATTTTATTATACACTT 813  
Qy 806 ValAsnAlaGln 809  
Db 814 GCTATGCAACAA 825  
RESULT 15  
BHL55700  
LOCUS ENTDS67R Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION genomic, DNA sequence.  
ACCESSION BHL55700  
VERSION BHL55700.1 GI:15727822  
BHL55700 913 bp DNA linear GSS 24-SEP-2001

## KEYWORDS

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica

REFERENCE Eukaryota; Entamoebidae; Entamoeba.

AUTHORS 1 (bases 1 to 913), Van Aken, S. and Fraser, C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HM1:IMSS sheared DNA library (2001)

COMMENT Unpublished (2001)

CONTACT: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjoftus@tigr.org

Clones are derived from the Entamoeba histolytica HM1:IMSS sheared

DNA library

Seq primer: M13-Reverse

Class: shotgun

High quality sequence start: 7

High quality sequence stop: 805.

Location/Qualifiers

1..913

## FEATURES

source

/organism="Entamoeba histolytica"

/strain="HM1:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/notes="Vector: pGEM-1; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broths cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 Kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

BASE COUNT 359 a 118 c 185 g 251 t

## ORIGIN

## Alignment Scores:

Pred. No.: 4.17e-72 Length: 913

Score: 711.50 Matches: 146

Percent Similarity: 69.37% Conservative: 42

Best Local Similarity: 53.87% Mismatches: 82

Query Match: 17.44% Indels: 2

DB: 17 Gaps: 1

US-09-873-409-2 (1-812) x BH155700 (1-913)

QY 533 TyrSerLysAlaLysSerGlyValAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 552

DB 29 TTTGCTACAGCTAAAGCTTCTGCATATAGATTATCAACAATGATAGATCCAGAT 88

QY 553 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 572

DB 89 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGAATGAAATATTACATTA 148

QY 573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592

DB 149 GAAGATGTTCAATTAGATATATCCAAACAGACCAACTAAACAAATCTTGTGGGACTTGAT 208

QY 593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612

DB 209 CTTGAATTAAGAAAGGACAAACAGTTCATTAGTAGGAGCATCAGGATGTGGTAAATCA 268

QY 613 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632

DB 269 ACTACTATTCAATTAGTCCAAAGAAATATTATGATCCAGTTGGTGGGTCAAGTCAAAATTAGAT 328

QY 633 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 652

DB 329 CGAAAAGATTAAAGAGATTAAATATCAATGCTTAAGAAATCAATAGGATTAGTTGGA 388

QY 653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaIleValPro 672

DB 389 CAAGAACCTTATTTGTCATGTACTATTAGAGAAATATTATGCTTGGAGCTAGAGAT 448

QY 673 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 691

DB 449 CGAAACACCAACTCAAGACAGAGATGATGATGCTTAAATGCAATGCAATGAA 508

QY 692 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 711

DB 509 TTTATTTCTCATCTTCCAGAGGATATGATACAAATGCTAGGAGAAAAGAGCTGCATTA 568

QY 712 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIle 731

DB 569 TCAGGAGGACAAACAAAGAAATTTGCTATTGACGCTGCTATTGATTAAGAAACCTCAAT 628

QY 732 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 751

DB 629 TTATTTACTTGTAGAGCTACATCAGCACTTGATACACAAGTGAAAAGATTGTACACAA 688

QY 752 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 771

DB 689 GCATTTGAAAAGCTAGTCAAGGAGAACAAACAATTTGTTGTAGCACATAGATTAACT 748

QY 772 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 791

DB 749 GTTAGAAATGCAAGTAGATAAATTTGTTATTCCTCATCAAGGAGAAATTTAT-TTGACAGGAACA 807

QY 792 HisGlnGluLeuLeuArgAsnArgAspIleTyr 802

DB 808 CATTCAAGATTAAATGATTTGAAGGGGACATAT 840

Search completed: March 31, 2003, 13:49:57

Job time : 5612.21 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 118.913 Seconds  
(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-2  
Perfect score: 4079  
Sequence: 1 MVDENDRALNVRHIDHG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09873409/runat\_27032003\_115422\_19259/app\_query.fasta\_1.7544  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1.578 @runat\_27032003\_115422\_19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2436	59.7	4264	2	US-08-784-649A-1
2	2436	59.7	4264	2	US-08-784-649A-5
3	2436	59.7	4646	1	US-08-181-471-2
4	2436	59.7	4669	6	5206352-3
5	2436	59.7	6505	2	US-08-793-610-5
6	2436	59.7	9318	2	US-08-793-610-6
7	2413.5	59.2	4233	3	US-09-120-513-1
8	2413.5	59.2	4233	4	US-09-450-105-1
9	2412	59.1	4669	2	US-08-752-447-1
10	2412	59.1	4669	4	US-09-316-167-1
11	2397	58.8	4669	2	US-08-583-276-18
12	2280	55.9	2726	1	US-08-461-823-1

13	1690.5	41.4	4002	2	US-08-996-545-1	Sequence 1, Appli
14	1690.5	41.4	4002	2	US-08-996-545-3	Sequence 3, Appli
15	1690.5	41.4	4002	4	US-09-328-320-1	Sequence 1, Appli
16	1690.5	41.4	4002	4	US-09-328-320-3	Sequence 3, Appli
17	1661	40.7	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1618	39.7	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1570.5	38.5	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	1474	36.1	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1255.5	30.8	3909	1	US-08-232-537-1	Sequence 1, Appli
22	1223.5	30.0	3924	2	US-08-996-644-3	Sequence 3, Appli
23	1223.5	30.0	3924	3	US-09-352-552-3	Sequence 3, Appli
24	1223.5	30.0	3927	2	US-08-996-644-1	Sequence 1, Appli
25	1223.5	30.0	3927	3	US-09-352-552-1	Sequence 1, Appli
26	1172	28.7	6143	1	US-08-612-521-3	Sequence 3, Appli
27	1116.5	27.4	3792	4	US-09-351-224E-10	Sequence 10, Appli
28	1032.5	25.3	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	915	22.4	13188	4	US-08-961-527-70	Sequence 70, Appli
30	824	20.2	6492	4	US-08-961-527-188	Sequence 188, App
31	793	19.4	7760	4	US-08-961-527-63	Sequence 63, Appli
32	776	19.0	7186	4	US-08-961-527-39	Sequence 39, Appli
33	763	18.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
34	763	18.7	4411529	4	US-09-103-840A-1	Sequence 1, Appli
35	678	16.6	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	644	15.8	2244	4	US-09-061-764A-18	Sequence 18, Appli
37	623.5	15.3	2061	4	US-09-061-764A-17	Sequence 17, Appli
38	608.5	14.9	4781	2	US-09-001-273-1	Sequence 1, Appli
39	608.5	14.9	4781	4	US-08-843-459A-1	Sequence 1, Appli
40	608.5	14.9	4847	3	US-09-061-400-1	Sequence 1, Appli
41	608.5	14.9	5120	3	US-08-772-270A-6	Sequence 6, Appli
42	608.5	14.9	8370	2	US-08-488-706-1	Sequence 1, Appli
43	601	14.7	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
44	600.5	14.7	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
45	594.5	14.6	7721	3	US-08-772-270A-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1  
US-08-784-649A-1  
; Sequence 1, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,649A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: Reg. No. 5830697 36,677  
; REFERENCE/DOCKET NUMBER: 060377/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5070  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4264 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-784-649A-1

```

**Alignment Scores:**

Pred. No.:	6,886-275	Length:	4264
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.77%	Indels:	14
DB:	2	Gaps:	4

US-09-873-409-2 (1-812) x US-08-784-649A-1 (1-4264)

Qy	2	ValAspGluAsnAspIleAlaArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal	21
Db	1495	GTTCATGGACAGATATTAGGACCAATAATGAAGGTTTCTACGGGAATCATTTGGTGTCG	1554
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg	41
Db	1555	GTGAGTCAGGAACCTGATTGTTGCCACACAGATAGTGAACAACATTCGCTATGGCGGT	1614
Qy	42	AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
Db	1615	GAAAAATGTCACCATGATGAGATTGAGAAGCTGTCAAGGAAGCCNAATGCCATATGACTTT	1674
Qy	62	IleMetClnPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1675	ATCATGAAATCTGCCTCATAAATTGACACCCCTGGTTGGAGAGAGGGGCCAGTTGAGT	1734
Qy	82	GlyClyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1735	GGTGGCAGAAGCAGAGGATCGCCATTGCATGCTGCCCTGGTTTCGCAACCCCAAGATCCTC	1794
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1795	CTGCTGGATGAGGCCACGTCAGCCTTTGGACACAGAAAGCGAAGCAGCTGGTTTCAGGTGCT	1854
Qy	122	LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle	141
Db	1855	CTGGATTAAGGCCAGAAAGAAGGTCCGACCACCATTTGTATGATCTCATCTGTTGCTACAGTT	1914
Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1915	CGTAATGCTCAGCTCATCGCTGGTTTCGGATGATGAGTCACTTGTGCAGAAAGGAATCAT	1974
Qy	162	AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln-----	178
Db	1975	GATGAATCTATGAAGAGAAAGACATTTACTTCAAACTTGTCACAATGCACACAGCAGGA	2034
Qy	179	-----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	195
Db	2035	AATGAAGTTCAATTAGAAAATCGACGTGATGAATCCAAAAGTGAAATTTGATCGCTTGAA	2094
Qy	196	ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle	209
Db	2095	ATGTCTTCAAATGATTCAAGATCCAGTCTAATTAAGAAAAAAGATCAACTCGTAGAGATGTC	2154
Qy	210	Lys--SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu	228
Db	2155	CGTGATCAACACCCACAGCAAGAGCTTAGTACCAAGAGGCTCTGGATGAAGTATA	2214
Qy	229	ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal	248
Db	2215	CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTATAATTAACCTGAATGCGCTTATTTGTT	2274
Qy	249	LeuCluYThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe	268
Db	2275	GTTCGGTGATATTTTGGCCATTTAAATGAGGCGCTCGCAACACAGCATTTTGGCAATTAATTT	2334
Qy	269	AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla	287

```
Db 3415 AAAGTCTGCTGATGGCAAGAAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGCACAC 3474
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3475 CTGGGCATCGTGTCCAGAGCCCATCTGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGT 3534
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleIleIleIleIleIleIleIle 687
Db 3535 TATGGAGACAACAGCCGGTGGTGTACAGGAAGAGATTTGTAGGGCAGCAAGAGAGGCC 3594
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluGlyTyrAsnThrGlnValGlyLeuLys 707
Db 3595 AACATACATCGCTTCATCGAGTCACTGCTGCTTAATAATATAGCACTAAGTAGGACAAA 3654
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaAlaAlaAlaAlaLeu 727
Db 3655 GGAACCTCAGCTCTCTGCTGGCCAGAAACACGATTCAGTGTGCTGCTGCTGCTGCTGCT 3714
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3715 CAGCCCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCGGATACAGAAAGTGAAGAG 3774
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3775 GTTGTCCAGAGCCCTGGACAAAGCCAGAGAGCGCGACCTGCAATTGTGATGCTCAC 3834
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3835 CGCCTGTCCACCATCCAGAAATGCAGACTTAATAGTGTGTGTTCCAGAAATGCAGAGTCAAG 3894
Qy 788 GluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3895 GAGCATGGCAGCATCAGCAGCTGTGGCAGACAGAAAGGATCTATTTTCAATGCTCAGT 3954
Qy 808 AlaGlnSer 810
Db 3955 GTCCAGGCT 3963

RESULT 2
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.: 6.88e-275 Length: 4264
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
Db: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-784-649A-5 (1-4264)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1495 GTTGATGGACAGATATTAGGACCAATAATGTAAGTTTCTACGGGAATCATTTGTGTG 1554
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1555 GTGAGTCAGAAACCTGATTGTTTGGCCACCACGATAGCTGAAAACATTGCTATGCGCGT 1614
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1615 GAAATATGTCAACATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1674
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81
Db 1675 ATCATGAACCTCCCTCATAAATTTGACACCCCTGGTGGAGAGAGGGGCCCATTTGAGT 1734
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1735 GTTGGCAGAAAGCAGAGGATCGCCATTGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1794
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1795 CTGCTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCAAGCAGTGGTTCAGGTGGCT 1854
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1855 CTGGATAAGCCAGAAAGGTCCGACCCATTGTTGATAGCTCATGCTTTGTCTACAGT 1914
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1915 CGTAATGCTCAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1974
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 1975 GATGAACCTCATGAAAGAGAAAGGCACTTTACTTCAAACTTGTCAAAATGCAGACAGCA 2034
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 2035 AATGAAGTTGAATAGAAATGCAGCTGATGATGATGATGATGATGATGATGATGATGAT 2094
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 2095 ATGCTCTCAATGATTCAAGATCCAGTCTTAATAAGAAAGATCAACTCGTAGGAGTGTC 2154
Qy 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 2155 CTGGATCACAAGCCCAACAGCAAGAGCTTAGTACCAGAGAGGCTCTGGATGAAGATATA 2214
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248
Db 2215 CCTCCAGTTTCTTTTGGCAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTTTGT 2274
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2275 GTTGGTGTATTTTGTGCCATTATAAATGAGGCGCTGCAACACGACATTTGCAATAATTT 2334
```

Qy	269	AlaIysIleIleThrMetPheGlyAen---AenAspLysThrThrLeuLysHisAspAla	287
Db	2335	TCAAAGATTATAGGGGTTTTTACAAAGATTGATGATCTGAAACAAACCGACAGAAATAGT	2394
Qy	288	GluIleTySerMetIlePheValIleLeuGlyValIleCysPheValSerTyrTyrPheMet	307
Db	2395	AAC TTGTTTTCATATGTTTCTAGCCCTTGGAAATATTTCTTTATTACATATTTTCCCT	2454
Qy	308	GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla	327
Db	2455	CAGGTTTCACATTGGCAAGCTGGAGAGATCTCTCCACCAAGCGCTCCGATACATGGTT	2514
Qy	328	PheIysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly	347
Db	2515	TTCCGATCCATGCTCAGACAGGATGTGATGTGGTTTGATGACCCCTCAAAAAACCCACTGGA	2574
Qy	348	GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg	367
Db	2575	GCATTGACTACCAAGCTCGCAATGATGCTCTCAAGTTTAAAGGGGCTATAGGTTCCAGG	2634
Qy	368	IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle	387
Db	2635	CTTGCTGTAATTACCCAGAAATAGCAAAATCTGGGACAGGAATAATATATCTTCATC	2694
Qy	388	TyrGlyTrpGluMetThrPheIleLeuSerIleAlaProValLeuAlaValThrGly	407
Db	2695	TATGGTTGGCAACTAACACACTGTTACTCTTAGCAATTGTACCCATCATTCATTAAGCAGGA	2754
Qy	408	MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis	427
Db	2755	GTGTTGTAATGAAATGTTGTCTGGCAAGCACCTGAAAGATAGAAAGAACTAGAAAGGT	2814
Qy	428	AlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIleValSerLeuThrArg	447
Db	2815	GCTGGGAGAGATCCTACTAGACAAATAGAAAACCTCCGAAACGTTGTTCTTTGACACTCAG	2874
Qy	448	GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer	467
Db	2875	GAGCAGAAGTTTGAACATATGATGCTCAGAGTTTTCAGGTACCATACAGAAACTCTTTG	2934
Qy	468	LysIysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla	487
Db	2935	AGGAAGACACACATCTTTGGAATTACATTTTCTCTCACCCAGGCACATGATGATTTTTC	2994
Qy	488	TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu	507
Db	2995	TATGCTGGATGTTTCGGTTTGAGCGCTACTTTGGTGGCACATAAATCATGAGCTTTGAG	3054
Qy	508	GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu	527
Db	3055	GATGTTCTGTAGTATTTTCAGCTGTGTCTTTGGTGCCATGCGCGTGGGCAAGTCAGT	3114
Qy	528	ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu	547
Db	3115	TCATTTGCTCCTGACTATGCCAAGCCAAAATATATCAGCAGGCCACATCATCATCATCAT	3174
Qy	548	GluLysIysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu	567
Db	3175	GAAAAACCCCTTTGATTGACACTACAGCAGGAAGGCCCTTAATGCGCAACACATTCGGAA	3234
Qy	568	GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle	587
Db	3235	GGAAATGTCATTTGGTGAAGTTGTTTCACTATCCACCCGACCGGACATCCCATG	3294
Qy	588	LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer	607
Db	3295	CTTCAGGACTGAGCCTGGAGGTGAAGAAAGCGCCAGACGCTGGCTCTGGTGGGACGAGT	3354
Qy	608	GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly	627
Db	3355	GGCTGTGGGAAGAGACAGATGGTCTCAGCTCTCTGGAGCGGTTCTACGACCCCTTGGCAGGG	3414
Qy	628	GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln	647

```

Db 3415 AAAGTGCTGCTTGAATGGCAAAAGAAATAAAGCAGCTGAATGTTTCAAGTGGCTCCAGCAGCAC 3474
Qy 648 ILeAlaileValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3475 CTGGGATCGTGTCCAGAGGCCATCTCTGTTGACATGCAGCATTCGTGAGACATTTGCC 3534
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluLeuIysGluAlaAlaAsnAlaAla 687
Db 3535 TATGGAGACAACAGCCGGTGGTGTACAGGAAGAGATTGTGAGGGCAGCAAAAGGAGGCC 3594
Qy 688 AsnIleHisSerPheIleGluClyLeuProGluIysTyrAsnThrGlnValGlyLeuIys 707
Db 3595 AACATACATGCTTCATCGAGTCATGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3654
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnIysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3655 GGAACCTAGCTCTCTGTGGCCAGAAACAACGATGCCATAGCTCTGTGCCCTGTGTAGA 3714
Qy 728 LysProIysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluIys 747
Db 3715 CAGCCTCATATTTTGTCTTTGGATCAAGCCACGTCAGCTCTGGGATACAGAAAGTGAAAG 3774
Qy 748 ValValGlnHisAlaLeuAspIysAlaGthrGlyArgThrCysLeuValValThrHis 767
Db 3775 GTTGTCCAGAAGCCCTGGACAAAGCCAGAGAGCGCCACCTCGATTTGTGATTGCTCAC 3834
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3835 CGCTGTCCACATCCAGATCCAGACTAATAGTGTGTGTTTCAGATGCCAGAGTCAAG 3894
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3895 GAGCATGGCAGCATCAGCAGCTGTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 3954
Qy 808 AlaGlnSer 810
Db 3955 GTCCAGGCT 3963

RESULT 3
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishto, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: AHT0029P
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 619-792-3680  
 ; TELEFAX: 619-792-8477  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4646 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 425..4267  
 ; US-08-181-471-2

Alignment Scores:  
 Pred. No.: 8e-275 Length: 4646  
 Score: 2436.00 Matches: 466  
 Percent Similarity: 76.79% Conservative: 166  
 Best Local Similarity: 56.62% Mismatches: 177  
 Query Match: 59.72% Indels: 14  
 DB: 1 Gaps: 4

US-09-873-409-2 (1-812) x US-08-181-471-2 (1-4646)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 DB 1781 GTTGATGGACAGATATTAGGACCATTAATGTAAAGTTTCTACGGGAATCATTTGGTGTG 1840  
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIysTyrGlyArg 41  
 DB 1841 GTGAGTCAGGAACCTGTAATTTGTCACACCATAGTGTGAAACATTCGCTATGCGCGT 1900  
 QY 42 AspAspValThrAspGluMetGluArgAlaAargGluAlaAsnAlaTyrAspPhe 61  
 DB 1901 GAAATGTCCACATGATGAGATGTGAAAGCTGTCAAGGAAGCCAAATGCTATGACTTT 1960  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 DB 1961 ATCATGAACCTGCTCATTAATTTTGACCCCTGTTGGAGAGAGAGGGCCCAAGTTGAGT 2020  
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAargAlaLeuValArgAsnProIysIleLeu 101  
 DB 2021 GGTGGGACAGACAGAGGATCGCATTCGACGTGCTGCTGTCGCAACCCCAAGATCCTC 2080  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAlaA 121  
 DB 2081 CTGCTGGATGAGGCCCGTCAGCTTTGGACACAGAAGCGAAGCGTTCAGGTGGCT 2140  
 QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
 DB 2141 CTGGATTAAGCCAGAAAGGTGGACACCAATTTGTATAGTCTATCGTTTGTCTACAGTT 2200  
 QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
 DB 2201 CGTAATGCTGACGTCTCGTGGTTTCGATGATGGAGTCATTCGAGAGAAGGAATCAT 2260  
 QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerIleuValMetSerGln----- 178  
 DB 2261 GATGAATCATGAAGAGAAAGGCATTTACTTCAAACTTTGTCAAAATGTCAGACAGCAGGA 2320  
 QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
 DB 2321 AATGAAGTTGAATTAGAAATGACGTGATGAATCCAAAGTCAAAATGATGATGCTTGGAA 2380  
 QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
 DB 2381 ATGCTCTTCAATGATTTCAAGATCCAGTCTAATTAAGAAAGATCAACTCGTAGGAGTGTC 2440  
 QY 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228  
 DB 2441 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAGAGAGGCTCTGGATGAAGATATA 2500

QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248  
 DB 2501 CTTCCAGTTTCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTGTTGT 2560  
 QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
 DB 2561 GTTGGTGTAATTTGTGCCATTATAAATGGAGGCTCCAAACCAAGCATTTGCAATAATATT 2620  
 QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287  
 DB 2621 TCNAAAGATTATAGGGGTTTTTACAGAATTGATGATCCTGAAACAAACACAGATAAGT 2680  
 QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
 DB 2681 AACTTGTCTTCACTATTGTTCTAGCCCTTGAATATTCTTTTATTACATTTTCTCTT 2740  
 QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
 DB 2741 CAGGGTTTCACATTTTGGCAAGCTGGAGAGATCCTCACCAGCGGCTCCGATACATGGTT 2800  
 QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
 DB 2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCTTAAACACCATCTGGA 2860  
 QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
 DB 2861 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGCTATAGTTCAGG 2920  
 QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
 DB 2921 CTGTGTTGTAATACCAGAAATATAGCAAAATCTTGGACAGAGATAAATATATATCTTCATC 2980  
 QY 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaproValLeuAlaValThrGly 407  
 DB 2981 TATGGTTGGCAATCAACTGTTACTCTTAGCAATTTGTACCCATCAATGCAATAGCAGGA 3040  
 QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
 DB 3041 GTTGTGAAATGAAATGTTGCTGACACAGCCTGAAAGATAAGAAAGAACTAGAAAGT 3100  
 QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAargThrIleValSerLeuThrArg 447  
 DB 3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAACTTCGAAACCGTTCTTTCTTTGACTCAG 3160  
 QY 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
 DB 3161 GACGAGAAGTTTGAACATATGATGCTCAGAGTTTCAGGTACCATACAGAAACTCTTTG 3220  
 QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
 DB 3221 AGGAAGACACATCTTTTGGAAATTACATTTTCTTCCACCAGGCAATGATGATTTTTC 3280  
 QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
 DB 3281 TATGCTGGATGTTTCCGGTTTGGAGCTACTTGGTGGCACATAAACTCATGAGCTTGG 3340  
 QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
 DB 3341 GATGTTCTGTAGTATTTTTCAGCTGTTGCTTTGGTGGCATGGCCGTGGGGGAAGTCAGT 3400  
 QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 547  
 DB 3401 TCATTTGCTCTGACTATTCGCAAGCCAAATATATCAGAGAGCCCATCATCATGATCAT 3460  
 QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
 DB 3461 GAAAAAACCCCTTTGATTGACAGCTACAGCAGCGAAGGCTTAATGCCGAACACATTTGGA 3520  
 QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
 DB 3521 GGAAATGTCACATTTTGGTGAAGTTGTTATTCAATATCCCAAGCCGACCGGACATCCCGAG 3580

```

Qy 588 LeuArgGlyLeuSerLeuSerilecluarGlyLysThrValAlaPheValGlySerSer 607
Db 3581 CTTGAGGACTGAGCTGGAGCTGAAGAAGGCCAGACGCTGGCTGGTGGCAGCAGT 3640

Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuThrAspProValGlnGly 627
Db 3641 GGCTGTGGGAAGACACAGTGGTCCAGCTCTGGAGCGGTCTACGACCCCTTGGCAGGG 3700

Qy 628 GlnValLeuPheAspGlyValAspAlaLysGlnLeuLeuValGlnTrpLeuArgSerGln 647
Db 3701 AAGTGTGCTGTGATGGCAAGAATAAAGCAGCTGAATGTTCAGTGGCTCCGAGCACAC 3760

Qy 648 IleAlaIleValProGlnGlnProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3761 CTGGGCATCGTGTCCAGAGGCCCATCTGTTGACTGCACATTCGCTGAGAACATTGCC 3820

Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3821 TATGGAGACAACAGCCGGTGTGTACAGGAAGAGATCGTGGAGGCGCAAGAGGCC 3880

Qy 688 AsnIleHisSerPheIleGluGlyLeuProGlnLysTyrAsnThrGlnValGlyLeuLys 707
Db 3881 AACATACATGCTTCATCGAGTCTGCTTAATAAATATAGCATTAAGTAGGAGACAAA 3940

Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3941 GGAACCTCAGCTCTCTGGTGGCCAGAAACAGCATTCGCTGCTGGCTGTTGTTAGA 4000

Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLys 747
Db 4001 CAGCCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGA 4060

Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4061 GTTGTCCAAAGCCCTCGGACAAAGCCAGAGAGCCGACCTGCATGTGATGCTCAC 4120

Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4121 CGCCTGTCCACCATCCAGATCCAGACTTAATAGTGGTGTTCAGAAATGGCAGAGTCAAG 4180

Qy 788 GluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleThrPheLysLeuValAsn 807
Db 4181 GAGCATGGCAGCATCAGCAGCTGTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 4240

Qy 808 AlaGlnSer 810
Db 4241 GTCCAGGCT 4249

RESULT 4
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3

Alignment Scores:
Pred. No.: 8,07e-275 Length: 4669
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177

```

```

Query Match: 59.72% Indels: 14
DB: 6 Gaps: 4
US-09-873-409-2 (1-812) x 5206352-3 (1-4669)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTGATGGACAGATATTAGGACCAATAATGTAAGGTTTCTACGGGAATCATATTGGTGTG 1840

Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGAACTGTATTGTTGGCCACCACCATAGCTGAAAACATTCGTATGGCCGT 1900

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCCATTCATGATGATGAGATGAGAAAGCTGTCAAGGAGCAATGCTATGACTTT 1960

Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAACTGCTCATATAAATTTGACACCTGTGTTGGAGAGAGAGGGCCAGTTGAGT 2020

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2021 GGTGGGAGAGAGAGAGATCGCCATTTGCACGTCCTGCTGTTGCGCAACCCCAAGATCCTC 2080

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 2081 CTGCTGATGAGGCCAGCTCAGCCTTGGACACAGAAAGCCAGAGCAGTGGTTCAGTGGCT 2140

Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 2141 CTGGAATAAGCCAGAAAGGTTCGACCCACCATTTGTGATAGTCTCATCTGTTGTCTACAGTT 2200

Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 2201 CGTAATGCTCAGCTCATCTGCTGCTGTTGATGATGAGTCAATTTGGAGAAGGAATCAT 2260

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 2261 GATGAATCATGAAGAGAAAGGCAATTTACTTCAACTTGTCAATGTCAGACAGACAGGA 2320

Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 2321 AATGAAGTTCAATTTAGAAAATGCAAGTGTGATGAATCCAAAAGTGAATGATGCTCGTGA 2380

Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 2381 ATGCTCTTCAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGT 2440

Qy 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
Db 2441 CGTGGATCACAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGATATA 2500

Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPropheValVal 248
Db 2501 CTTCCAGTTTCCTTTGGAGGATTAAGAGCTAAATTTAACTGAATGCTTATTTGTT 2560

Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCTGCAACACAGCATTTGCAATAATATT 2620

Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2621 TCAAGATTAATAGGGTTTTTACAGAAATGTATGATCTCTCAAAACAAACACAGCAATAGT 2680

Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2681 AACTGTGTTTCACTATTGTTTCTAGCCCTTGGAAATATTATTTCTTTTATTACATTTTCT 2740

Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2741 CAGGGTTTCATATTGGCAAAAGTGGAGAGATCTCTCAACCAAGCGGCTCCGATACATGTT 2800

```



Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347  
Db 2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGAGCCCTAAAAACACCATGGA 2860  
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
Db 2861 GCATTGACTACAGGCTCGCAATGATCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2920  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 2971 CTTGCTGTAATCCAGAAATATAGCAATCTTGGACAGGAATAATATATATCTTCATC 2980  
Qy 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2981 TATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCCATCATTTGCAATGACGGA 3040  
Qy 408 MetIleGluThrAlaAlaMetThrClyPheAlaAsnLysAspLysGlnIleuLysHis 427  
Db 3041 GTTGTGTAATGAAATGTTGTGGCAAGCCTGGAAGAGATAGAAAGAACTAGAAGGT 3100  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
Db 3101 CTTGGAGAGATCGCTACTGAGCAATAGAAACTTCCGAAACCGTTGTTCTTGACTCAG 3160  
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
Db 3161 GAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220  
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 3221 AGGAAAGCACACATCTTTGGAAATATACATTTCTTCCTCCACCGCAATGATGATATTTTCC 3280  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 3281 TATGCTGATGTTTCGGTTTGGAGCCTACTTGGTGGCACATAACTCATGAGCTTTGAG 3340  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
Db 3341 GATGTTCTGTTAGTATTTTCAGCTTCTGCTTGTGTCATGGCCGTCGGGCAAGTCAGT 3400  
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
Db 3401 TCATTGCTCTGCTGATATGCCAAAGCCAAAATATATCAGCAGGCCCATCATCATGATCAT 3460  
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
Db 3461 GAAAGAAACCCCTTGTATGACAGCTACAGCAGGAGGCTTAATCCGAACACATTGGA 3520  
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
Db 3521 GGAATATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCACCGGACACATCCAGTG 3580  
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607  
Db 3581 CTTCAGGAGCTGAGCTGAGGTGAAGAGGCGCACAGCTGCTGCTGGTGGCAGCAGT 3640  
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627  
Db 3641 GCCTGTGGAGAGACACAGTGTGTCAGCTCTCTGGAGCGGTTCTACAGCCCTTGGCAGGG 3700  
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647  
Db 3701 AAAGTGTCTTGTATGGCAAGAAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGCACAC 3760  
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667  
Db 3761 CTGGGCATGCTGCTCCAGGAGCCCATCTGTTGATGCTGAGCAATTTGCTGAGAACATGCC 3820  
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687  
Db 3821 TATGAGACAAACAGCCGGTGTGTGCACAGAAAGAGATCGTGAGGGCAGCAAGAGGCC 3880  
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707

Db 3881 AACATACATGCTTTCATCGAGTCACTGCCCTAAATAATATAGCACTAAAGTAGGACAAA 3940  
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727  
Db 3941 GGAACCTCAGCTCTCTGTTGGCCAGAAACACGCAATTCGCATAGCTCGTCCCTTGTAGA 4000  
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747  
Db 4001 CAGCCTCATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGA 4060  
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
Db 4061 GTTGTCCAGAGCCCTGGGACAAAGCCAGAGAGCCGACCTGCATTTGATTTGCTCAC 4120  
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787  
Db 4121 CGCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGATGGCAGAGTCAAG 4180  
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807  
Db 4181 GAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 4240  
Qy 808 AlaGlnSer 810  
Db 4241 GTCCAGGCT 4249  
RESULT 5  
US-08-793-610-5  
; Sequence 5, Application US/08793610  
; Patent No. 5858744  
; GENERAL INFORMATION:  
; APPLICANT: BAUM, Christopher  
; APPLICANT: STOCKING-HARBERS, Carol  
; APPLICANT: OSTERTAG, Wolfram  
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,610  
; FILING DATE: 07-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 31 973.8  
; FILING DATE: 08-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 03 952.1  
; FILING DATE: 07-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/03175  
; FILING DATE: 10-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berman, Richard J.  
; REGISTRATION NUMBER: 39,105  
; REFERENCE/DOCKET NUMBER: FI614-7007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6505 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA  
US-08-793-610-5

## Alignment Scores:

Pred. No.: 1.45e-274 Length: 6505  
Score: 2436.00 Matches: 466  
Percent Similarity: 76.79% Conservative: 166  
Best Local Similarity: 56.62% Mismatches: 177  
Query Match: 59.72% Indels: 14  
DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-793-610-5 (1-6505)

QY 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21  
DB 3173 GTTGATGGACAGATATTAGGACCAATAATGTAAAGTTCTACGGGAATCATTTGGTGG 3232  
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
DB 3233 GTGAGTCAGAACTGTATTGTTGGCCACCATAGCTGMAAACATTCGCTATGGCCGT 3292  
QY 42 AspAspValThrAspGluGluMetGluArgAlaAalArgGluAlaAsnAlaTyrAspPhe 61  
DB 3293 GAAATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAGCCCAATGCCATGACTTT 3352  
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
DB 3353 ATCATGAAACTGCTCATAAATTTTGACCTGTGGAGAGAGGGCCCAATGTGAGT 3412  
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
DB 3413 GGTGGGCAGAAAGCAGAGGATCGCATTTGCAGTGCCTGTTGCAACCCCAAGATCCTC 3472  
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
DB 3473 CTGCTGGATGAGGCCACGTCAGCTTTGGACACAGAAAGCGAAGCAGTGTTCAGTGGCT 3532  
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
DB 3533 CTGGATAGGCCAGAAAGTCGACACCATTTGTAGTACATCATGTTTGTCTACAGTT 3592  
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
DB 3593 CGTAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 3652  
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
DB 3653 GATGAATCATGAAGAGAAAGCATTACTTCAAACTTGTCAAAATGCAGACAGCAGGA 3712  
QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
DB 3713 AATGAAGTTGAATTAGAAATGACGCTGATGAATCAAAAGTGCAAAATGATGCTTGGAA 3772  
QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
DB 3773 ATGCTCTCAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGC 3832  
QY 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228  
DB 3833 CGTGGATCAACAGCCCAAGACAGAAAGCTTAGTACCAAGAGAGCTCTGGATGAAGTATA 3892  
QY 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248  
DB 3893 CTTCCAGTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCCCTTATTGTT 3952  
QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
DB 3953 GTTGGGTGATTTGTGCCATTATAAATGGAGGCTGCAACCCAGCATTTGCAATAATATT 4012  
QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287  
DB 4012

DB 4013 TCAAGATTATAGGGGTTTTTACAAGAATTGATGATCTCTGAACAAAAACGACAGATAGT 4072  
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
DB 4073 AACTTTGTTTCTACTATTGTTCTAGCCCTTGGAATATTCTTTTATTACATTTTCCCT 4132  
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
DB 4133 CAGGGTTTTCACATTTGGCAAGCTGGAGAGATCCCTCACCAGCGGCTCGATACATGTT 4192  
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
DB 4193 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTGTATGACCCCTAAAAACACCACTGA 4252  
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
DB 4253 GCATTGACTACGAGCTGCCCAATGATGCTGCTCAAGTTAAAGGGCTATAGTTCCAG 4312  
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
DB 4313 CTTGCTGTAATTAACCCAGAAATATAGCAAAATCTTGGGACAGGAATAATTATATCCTTCATC 4372  
QY 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
DB 4373 TATGGTTGGCAATTAACACTGTTACTCTTAGCAATTTGACCATCATTCGAATAGCAGA 4432  
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHis 427  
DB 4433 GTTGTGAAATGAAATGTTGCTGGCAAGCCTCTGAAAGATAAGAAAGAACTAGAGGT 4492  
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
DB 4493 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTCTTGTGACTCAG 4552  
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
DB 4553 GAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACCTTTTG 4612  
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
DB 4613 AGGAAGACACACATCTTTGGAATTAATTTCTTCCACCCAGCAATGATGATTTTTC 4672  
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
DB 4673 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTGG 4732  
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
DB 4733 GATGTTCTGTTAGTATTTTCAGCTGTGTTGTTGTCCTATGCGCGGCGCAAGTCAGT 4792  
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
DB 4793 TCATTGCTCTGACTATGCCAAAGCCAAATAATCAGCAGCCCATCATCATGATCATTT 4852  
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPheProAspThrCysGlu 567  
DB 4853 GAAAAAACCCCTTGTATTGACAGCTACAGCACGGAAGGCTTAATGCCGCAACACATTGGAA 4912  
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
DB 4913 GGAAATGTCAATTTGTTGTAAGTTGTTATTAATATCCACCCCGGACATCCAGATG 4972  
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607  
DB 4973 CTTCAGGACTGAGCTGGAGGTGAAGAGGCGCAGACGCTGCTCTGTTGGGCGACGT 5032  
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627  
DB 5033 GGCTGTGGGAAGACACAGTGTGTCAGCTCTCGAGCGGTTCTACGACCCCTTGGCAGGG 5092  
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647  
DB 5093 AAAGTGCTCTCTTGGGCAAGAAATAAAGCGACATGAATGTTTCAGTGGTCCGACGACAC 5152

```
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 5153 CTGGGCATCGTCCAGGAGCCCATCTGTTGACTGCAGCATTTGCTGAGAACATTGCC 5212
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 687
Db 5213 TATGAGACAACAGCCGGTGGTGTCAACAGGAAGAGATCGTGAGGCGACCAAGAGGCC 5272
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 5273 AACATACATGCTTCATCGAGTCATCTGCTTAATAATATAGCACTAAAGTAGGACACAA 5332
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 5333 GGAACCTCAGTCTCTGTTGGTGGCCAGAAACAAACGCAATGCGCTCGTGGCCCTTGTAGA 5392
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLys 747
Db 5393 CAGCCTCATATTTTCTGTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 5452
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 5453 GTTGTCGAAGAGCCCTGGACAAACCCAGAGAGAGCCGCGACCTGCTGATGTGCTCTCAC 5512
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 5513 CGCTGTCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAGAAATGCCAGAGTCAAG 5572
Qy 788 GluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 5573 GAGCATGGCAGCCATCAGCAGCTGCTGGCAGAGAGGATCTATTTTCAATGCTCAGT 5632
Qy 808 AlaGlnSer 810
Db 5633 GTCCAGGCT 5641

RESULT 6
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; US-08-793-610-6

Alignment Scores:
Pred. No.: 2,72e-274 Length: 9318
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 2 Gaps: 4
```

US-09-873-409-2 (1-812) x US-08-793-610-6 (1-9318)

```
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 3132 GTTGATGACAGGATATTAGGACCATTAATGTAAGTTTCTACGGGAAATCATTTGGTGTG 3191
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 3192 GTGAGTCAGGAACCTGTATTGTTTGCACACCATGATAGCTGAAACATTCGCTATGCCGCT 3251
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 3252 GAAATGTCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCATGCCATGACCTTT 3311
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 3312 ATCATGAAATGCTCTATAAATTTGACACCTGTTGGAGAGAGAGGGGCCAGTTGAGT 3371
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 3372 GGTGGGCAAGACAGAGGATGCCATTCACGTGCCCTCGCTGGTTCGCAACCCCAAGATCCTC 3431
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 121
Db 3432 CTGCTGGATGAGCCACGTCAGCTTGGACACAGAAACGAGCAGTGGTTCAGGTGGCT 3491
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 3492 CTGGATAAGGCCAGAAAGGTCGGACCATTTGTGATAGCTCATCGTTTGTCTACAGTT 3551
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 3552 CGTAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGTGGAGAAAGAAATCAT 3611
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 3612 GATGAACATCATGAAGAGAAAGAGGCAATTTACTTCAAACCTGTGCACAAATGCAGACAGGA 3671
Qy 179 -----AspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 3672 AATGAAGTTGAATTAGAAATGCAGCTGATGAATCAAAAGTGAATGATGCTTGGAA 3731
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 3732 ATGCTCTCAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTG 3791
Qy 210 Lys----SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 3792 CGTGGATCAACAGCCCAAGAGAGAGGCTTAGTACCAAGAGGCTCTCGATGAAAGTATA 3851
```









Db 2210 TTTTCACAGATTGTAGGGGTTTTTTTCAAGAGACAGGACCATTGAACCAAAACCAACCGAAT 2269  
Qy 287 AlaGluileTyrSerMetilePheValilleLeuGlyValilleCysPheValserTyrPhe 306  
Db 2270 TGTAACTTGTTCCTTCCTCTCTTTCTGGTCATGGGAATGATTTCTTTTGTACGTACTTC 2329  
Qy 307 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluileLeuThrMetArgLeuArgHisLeu 326  
Db 2330 TTTCAGGCTTCACATTTTGGCAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATG 2389  
Qy 327 AlaPheLysAlaMetLeuTyrGlnAspilleAlaTyrPheAspGluLysGluAsnSerThr 346  
Db 2390 GTCTTCAAAATCCATCTCGCAGCAGATATAGCTGTTTGTATGACCAATAAAACCACT 2449  
Qy 347 GlyGlyLeuThrThreileLeuAlaileAspilleAlaGlnileGlnGlyAlaThrGlySer 366  
Db 2450 GGCTCGCTGACTACCAAGGCTCGTAGTACGCTTCTAATGTTTAAAGGGGCTATGGGCTCC 2509  
Qy 367 ArgilleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValilleleSerPhe 386  
Db 2510 AGGCTTGCTGTATTACCCAGAAATGTAGCAAACTTGGCACAGGAATATATCTTATCCTTA 2569  
Qy 387 IleTyrGlyTyrGluMetPheLeuileLeuSerileAlaProValleuAlaValThr 406  
Db 2570 GTCTATGGCTGGCAGCTTACATTTTACTTGTAGTAATATATACCACTCAATTCCTTGGGT 2629  
Qy 407 GlyMetileGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 426  
Db 2630 GGAATATTGAAATGAACTGTGTCTGCTCAAGCTTGAAGGACAAAGAAAGCTAGAG 2689  
Qy 427 HisAlaGlyLysileAlaThrGluAlaLeuGluAsnileAargThrIleValserLeuThr 446  
Db 2690 ATCTCTGGGAAGATCGCTACAGAAGCAATGTGAAATCTTCGCACTGTTGTCTCTTTGACT 2749  
Qy 447 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 466  
Db 2750 AGGGAGCAGAAGTTTGAAACTATGATGCCCCAGAGCTTCAGATACCATACAGAAATGCT 2809  
Qy 467 SerLysLysAlaGlnileileGlySerCysTyrAlaPheSerHisAlaPheileTyrPhe 486  
Db 2810 TTGAAGAAGACACACGTCTTTGGGATCACCTTCGCTTCACCCAGGCCATGATTTATTTT 2869  
Qy 487 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuileGlnAlaGlyArgMetThrPro 506  
Db 2870 TCCTATGCTGCTTGTTCGGGTTCGGTGTCTACTTGGTGGCCAGAGAACTCATGACGTTT 2929  
Qy 507 GluGlyMetPheileValPheThrAlaileAlaTyrGlyAlaMetAlaileGlyLysThr 526  
Db 2930 GAAATGTTATGTGTGTAATTTCTGCTGTGTGTTTGGTGCCATGGCAGCAGGAAATACC 2989  
Qy 527 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 546  
Db 2990 AGTTCAATTCCTCTGACTACGGGAAGGCCAAAGTCTCAGCATCCCAATCATCAGGATC 3049  
Qy 547 LeuGluLysLysProAsnileAspSerArgSerGlnGluGlyLysLysProAspThrCys 566  
Db 3050 ATTGAGAAAATCCCCAGATTGACACGTACAGCAGCGAGGCTTGAAGCCTAATTCGTTA 3109  
Qy 567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586  
Db 3110 GAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCAACCCAGCCCAACATCCA 3169  
Qy 587 IleLeuArgGlyLysSerLeuSerileGluArgGlyLysThrValAlaPheValGlySer 606  
Db 3170 GTGCTTCAGGACTGAGCTTCGAGGTGGAAGAGGGGCAACGTTCCGCTGTGGGCAGC 3229  
Qy 607 SerGlyCysGlyLysSerThrserValGlnLeuLeuGlnArgLeuTyrAspProValGln 626  
Db 3230 AGTGGCTCGGGAAGAGTACAGTGGTCCAGCTGCTCGAGCGCTTCTACAACCCCATGGCT 3289  
Qy 627 GlyGlnValLeuPheAspGlyValasAlaLysGluLeuAsnValGlnThrLeuArgSer 646  
Db 3290 GGAACAGTGTCTTAGTAGGCAAGAAATAAAACAACTCAACGCTCAGTGGCTCCGCGCC 3349

Qy 647 GlnIleAlaileValProGlnGluProValleuPheAsnCysSerileAlaGluAsnile 666  
Db 3350 CACCTGGCATTTGTTCAGGAGCCCATCTGTTTGAATCAGCATCACCGAGAACATC 3409  
Qy 667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluileLysGluAlaAlaAsnAla 686  
Db 3410 GCCTACGGAGAACACAGCCGTGTCTGTCTCATGAGGAGATCGTGAGGCGCCCGCAGGAG 3469  
Qy 687 AlaAsnileHisSerPheileGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706  
Db 3470 GCCAACATCCACCACTTCATCGACTCACTGCTCGAGAAATACAAACACAGAGTGGGAGAC 3529  
Qy 707 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaileAlaArgAlaLeuLeu 726  
Db 3530 AAAGGCACTAGCTGTTCGGGGGGCAGGAAGCAGCGCATCGCGCGCCCTCGTC 3589  
Qy 727 GlnLysProLysileLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 746  
Db 3590 AGACAGCTTCATCTTACTTCTGATGAAGCGACATCAGCTCTGGATACGGAGAGTGA 3649  
Qy 747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766  
Db 3650 AAGTCTGTCAGGAAGCGCTGGACAAACCCAGGAAGCCGACCTGCTGATTCGGTCCGG 3709  
Qy 767 HisArgLeuSerAlaileGlnAsnAlaAspLeuileValLeuLeuHisAsnGlyLysile 786  
Db 3710 CACCGCTGTCCACCATCCAGAACCGACAGATGATCGTGGTGATTCAGAACCGCCAGGTC 3769  
Qy 787 LysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 806  
Db 3770 AAGGAGCAGCGGCCACCCAGCAGCGCTGTGCCCCAAGAGGCATCTATTTCTCGATGGT 3829  
Qy 807 AsnAla 808  
Db 3830 CAGGCT 3835  
RESULT 9  
US-08-752-447-1  
; Sequence 1, Application US/08752447  
; Patent No. 5994088  
; GENERAL INFORMATION:  
; APPLICANT: Mechtner, Eugene  
; TITLE OF INVENTION: Methods and Reagents for Preparing and  
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.  
; STREET: 300 South Wacker Drive, Seventh Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,447  
; FILING DATE: 15-NOV-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5994088nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 95,1121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-9808  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4669 base pairs



; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: 5'UTR  
 ; LOCATION: 1..424  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 425..4264  
 ; FEATURE:  
 ; NAME/KEY: 3'UTR  
 ; LOCATION: 4265..4669  
 ; US-08-752-447-1

## Alignment Scores:

Pred. No.: 5,29e-272 Length: 4669  
 Score: 2412.00 Matches: 462  
 Percent Similarity: 76.06% Conservative: 164  
 Best Local Similarity: 56.14% Mismatches: 183  
 Query Match: 59.13% Indels: 14  
 DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-752-447-1 (1-4669)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1781 GTTGATGGACAGCATATTAGGACCATAAATGTAAGGTTTCTACGGGAATCATTTGGTGTG 1840  
 QY 22 ValSerGlnProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1841 GTGAGTCAGGAACCTGTATTGTTTGGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900  
 QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaValAspPhe 61  
 Db 1901 GAAATGTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1960  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1961 ATCATGAACCTGCTCATAAATTTGCACCCCTGTTGGAGAGAGAGGGGCCAGTTGAGT 2020  
 QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 2021 GGTGGGAGAGAGAGAGATCGCCATTTGCACGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2080  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 2081 CTGCTGATGAGGCCAGCTGAGCCTTGAGACAGAAAGCGACAGTGGTTTCAGGTGGCT 2140  
 QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
 Db 2141 CTGGATAAGCCAGAAAAGGTGCGACCCATTTGTGATAGCTCATCTGTTTGTCTACAGTT 2200  
 QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
 Db 2201 CGTAATGCTCAGCTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2260  
 QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
 Db 2261 GATGAACCTCATGAAGAAGAAAGGCAATTTACTTCAAACTTGTCTCAATGCGACAGCAGGA 2320  
 QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
 Db 2321 AATGAAGTTGCAATTAGAAATGCGAGCTGATGATGATGATGATGATGATGATGATGATGAT 2380  
 QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
 Db 2381 ATGCTCTCAAATGATTCACAGATCCAGTCTAATAAGAAAAAAGATCAATCGTAGGAGTGC 2440  
 QY 210 LysSerAspPheIleAspLysAlaGluGluSerThr--GlnSerLysGluIleSerLeu 228  
 Db 2441 CGTGGATCACAGCCCNACACAGAAAGCTTAGTACCCAAAGAGGCTCTGGATGAAGATATA 2500

QY 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluThrProPheValVal 248  
 Db 2501 CCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTTATTTGTT 2560  
 QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
 Db 2561 GTTGGTGATTTTGTGCCATTATAAATGAGCCCTGCACACAGCATTTTCAATAATATT 2620  
 QY 269 AlaLysIleIleThrMetPheGlyAsn--AsnAspLysThrThrLeuLysHisAspAla 287  
 Db 2621 TCAGAATATTATAGGGGTTTTTACAGAATTGATGATCTCTGAAACAAACACAGATAGT 2680  
 QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
 Db 2681 AACTTGTCTTCTACTATTGTTCTAGCCCTTGGAAATATTCTTTATTATCAATTTTCTCT 2740  
 QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
 Db 2741 CAGGGTTTCACATTTGGCAAGCTGGAGAGATCTCACCAGCGGCTCCGATACATAGTT 2800  
 QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347  
 Db 2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGTTTCATGACCTAAACACCACTGGA 2860  
 QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
 Db 2861 GCATTGACTACCCAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAG 2920  
 QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
 Db 2921 CTTGCTGTAATTTACCAGAAATATAGCAATCTTGGACAGGAATTAATATATATCTTCATC 2980  
 QY 388 TyrGlyThrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
 Db 2981 TATGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTCGATCAGGAG 3040  
 QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
 Db 3041 GTTGTGTAATGAAATGTTTGTCTGCAAGCACTGAAAGATAAGAAAGAACTAGAAAGT 3100  
 QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
 Db 3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAATCTCCGAACCCGTTGTTCTTTGATCTAG 3160  
 QY 448 GlyLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
 Db 3161 GAGCAGAACTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAACTCTTTG 3220  
 QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
 Db 3221 AGGAAAGCACACATCTTTGGAATTTACATTTTCTTCCACCCAGGCAATGATGATTTTCTC 3280  
 QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
 Db 3281 TATGCTGGAGTTTTCGGTTTGGAGCTTACTTGTGTGGCAATATAAATCATGAGCTTTGAG 3340  
 QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeu 527  
 Db 3341 GATGTTCTGTAGTATTTTTCAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3400  
 QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
 Db 3401 TCATTGCTCTGCTGACTATGCCAAAGCCAAATATCAGACGCCCATCATCATGATCAT 3460  
 QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
 Db 3461 GAAAAACCCCTTGTATTGACAGCTACAGCAGGAAGGCTTAATGCCCGAACACATTGGAA 3520  
 QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
 Db 3521 GGAAATGTCACATTTGGTGAAGTTGTTATTCAACTATCCACCCGACCGACATCCAGTG 3580  
 QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607

Db	3581	CTTCAGGAGCTGAGCCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT	3640
Qy	608	GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly	627
Db	3641	GGCTGTGGGAAGACACACGTGGTCCAGCTCTGGAGCGGTGTACAGACCCCTTGGCAGGG	3700
Qy	628	GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln	647
Db	3701	AAAGTGTCTGTGATGGTCAAGAAATAAAGCGACTGAATGTTCACTGGCTCCGAGCACAC	3760
Qy	648	IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla	667
Db	3761	CTGGGCACTGTGCTCCAGGAGCCCATCTGTGTACATGCAGCATTTGCTGAGAACATTGCC	3820
Qy	668	TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla	687
Db	3821	TATGGAGACAAACAGCCGGTGGTCTCAGAGAGAGATCTGGAGGCGAGCAAGGAGGCC	3880
Qy	688	AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys	707
Db	3881	AACATACATGCTTTCATCGAGTCACTGCTCTAAATAAATATAGCACATAAAGTAGGAGACAA	3940
Qy	708	GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln	727
Db	3941	GGAACTCAGCTCTCTGGTGGCCAGAAACAGCATTCGCCATAGCTCGCTCGCTTGTGTAGA	4000
Qy	728	LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys	747
Db	4001	CAGCCTCATATTTTGTCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG	4060
Qy	748	ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis	767
Db	4061	GTGTCCAGAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCTGATGTGATGCTCAC	4120
Qy	768	ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys	787
Db	4121	CGCCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTTCAGAAATGGCAGAGTCAAG	4180
Qy	788	GluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn	807
Db	4181	GAGCATGCAGCATCAGCAGCTGCTGCGACAGAAAGGCATCTATTTTTCATGTCAGT	4240
Qy	808	AlaGlnSer 810	
Db	4241	GTCCAGGCT 4249	
RESULT 10			
US-09-316-167-1			
; Sequence 1, Application US/09316167			
; Patent No. 6365357			
; GENERAL INFORMATION:			
; APPLICANT: Mechethner, Eugene			
; APPLICANT: Roninson, Igor B			
; TITLE OF INVENTION: Methods and Reagents for Preparing and			
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein			
; NUMBER OF SEQUENCES: 2			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: McDonnell Boehrnen Hulbert & Berghoff Ltd.			
; STREET: 300 South Wacker Drive, Seventh Floor			
; CITY: Chicago			
; STATE: Illinois			
; COUNTRY: USA			
; ZIP: 60606			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/316,167			
; FILING DATE:			
; CLASSIFICATION:			

QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
DB 2321 AATGAAGTTGAATAGAAAATGACGCTGATGAATCCAAAAGTCAAAATTCATCGCTGGAA 2380  
QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
DB 2381 ATGCTCTCAATGATTCAAGATCCAGTCAATTAAGAAAAGATCAACTCGTAGGAGTGC 2440  
QY 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228  
DB 2441 CGTGGATCAACAGCCCAACACAGAAAGCTAGTACCAAGAGGCTCTGGATGAAGTATA 2500  
QY 229 ProGluValSerLeuLeuLysLysLeuLysLeuAsnLysProGluTrpProPheValVal 248  
DB 2501 CCTCCAGTTTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTTTGT 2560  
QY 249 LeuGlyThrLeuAlaSerValLeuLeuAsnGlyThrValHisProValPheSerIlePhe 268  
DB 2561 GTTGGTGTATTTTGTCCCAATTATAAATGGAGGCTCGAACCCAGCATTTGCAATAATATTT 2620  
QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287  
DB 2621 TCNAGATTATAGGGGTTTTTACAGAAATGATGATCTCGAAACAAACACAGAGATAGT 2680  
QY 288 GluIleThrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
DB 2681 AACTCTGTTTCACTATGTTTCTAGCCCTTGAATATTTCTTTTATTACATTTTCTTT 2740  
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
DB 2741 CAGGGTTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGGGCTCCGATACATGTTT 2800  
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaIleAlaTrpPheAspGluLysGluAsnSerThrGly 347  
DB 2801 TTCCGATCAGTCTCAGACAGAGATGAGTGGTTTCTATGACCTTAAACACACCATGGA 2860  
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
DB 2861 GCATTGACTACCAAGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGTTCAGG 2920  
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
DB 2921 CTTCGCTGTAATACCCAGAAATATAGCAAAATCTTGGACAGGAATAATATATCCTTCATC 2980  
QY 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
DB 2981 TATGGTTGGCACTAACACTGTTACTTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGA 3040  
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
DB 3041 GTTGTGTAATGAAATGTTTGTCTGGACAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3100  
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
DB 3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGAAACCGCTGTTCTTTTGAATCAG 3160  
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
DB 3161 GAGCAAGATTGAAACATATGATGATGCTCAGAGTTTGAGGTACCATACAGAAACTCTTTG 3220  
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
DB 3221 AGGAAGCACACATCTTTGGAATATACATTTCTTCACCCAGGCAATGATGATTTTTC 3280  
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
DB 3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACAATAAACTCATGAGCTTTGAG 3340  
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
DB 3341 GATGTTCTGTATGATTTTTCAGCTGTGTTCTTTGGTGGCCATGGCCGTGGGGCAAGTCAGT 3400

QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
DB 3401 TCATTTGCTCTGACTATGTCACAAAGATAATATCAGCAGCCCATCATCATCATGATCAT 3460  
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
DB 3461 GAAATAAACCCTTTGATTGACAGCTACAGCAGGAGGCTTAATCCGCAACACATTGGAA 3520  
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
DB 3521 GGAAATGTACATTTTGGTGAAGTTGTAATTAACATATCCACCCGACCGACATCCAGTG 3580  
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607  
DB 3581 CTTCCAGGACTGAGGCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGCCAGCAGT 3640  
QY 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627  
DB 3641 GGCTGTGGAGAGACACAGTGGTCCAGCTCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700  
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647  
DB 3701 AAGGTGCTGTTGATGGCAAGAAATAAAGCGACTGAATGTTCACTGGCTCCGACACAC 3760  
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667  
DB 3761 CTGGGCTGCTGCTCCAGGAGCCCATCTCTGTTGATGCTGAGCATTTGCTGAGAACATTGCC 3820  
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687  
DB 3821 TATGAGACAAACAGCCGGTGGTGTCTCAGGAAGAGATCTGAGGGCAGCAAGAGGCC 3880  
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707  
DB 3881 AACATACATGCTTCTCATCGAGTCACTGCTTAATAAATATAGCACTAAAGTAGCAGACAAA 3940  
QY 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727  
DB 3941 GGAACCTCAGCTCTCTGGTGGCCAGAAACAAACGCAATGGCATAGCTGCTGCTGCTTAGA 4000  
QY 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747  
DB 4001 CAGCTCATATTTCTTCTTTGATGAGCCAGCTCAGCTCTGATACAGAAAGTAGAAAG 4060  
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
DB 4061 GTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGCGCACCTGCTGATTTGCTCAC 4120  
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787  
DB 4121 CGCTGTCTCCACCATCCAGAAATCAGACTTAATAGTGGTGTTCAGAAATGGCAGAGTCAAG 4180  
QY 788 GluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807  
DB 4181 GAGCATGGCAGCATCAGCAGCTGCTGGCAGAAAGGCAATCTATTTTCAATGTCAGT 4240  
QY 808 AlaGlnSer 810  
DB 4241 GTCCAGGCT 4249  
RESULT 11  
US-08-583-276-18  
; Sequence 18, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:  
; APPLICANT: McDonagh, Kevin T.  
; APPLICANT: Nienhuis, Arthur  
; APPLICANT: Tolstoshev, Paul  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:

ADDRESS: Carella, Byrne, Bain, Gilfillan,  
ADDRESS: Cecchi & Stewart  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DW4.V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583.276  
FILING DATE: 05-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/332,444  
FILING DATE: 31-OCT-1994  
APPLICATION NUMBER: 07/887,712  
FILING DATE: 22-MAY-1992  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4669 bases  
TYPE: nucleic acid  
STRANDEDNESS: singular  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: Genomic DNA  
US-08-583-276-18

## Alignment Scores:

Pred. No.: 3,046-270 Length: 4669  
Score: 2397.00 Matches: 460  
Percent Similarity: 76.03% Conservative: 165  
Best Local Similarity: 55.96% Mismatches: 183  
Query Match: 58.76% Indels: 14  
DB: 2 Gaps: 4

US-09-873-409-2 (1-812) x US-08-583-276-18 (1-4669)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
Db 1781 GTTCATGGACAGATATTAGGACCAATAATGTAAGGTTCTACGGGAAATCATTTGGTGTG 1840  
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
Db 1841 GTGAGTCAGGAACCTGTATTGTTGCCACACGATAGCTGAAACAACTTCGCTATGGCCGT 1900  
Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
Db 1901 GAAATGTCACCATGATGAGATTGAGAAAGCTGTCAAGNAGCCATGCCATGACTTT 1960  
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
Db 1961 ATCATGAAATCGCTCATAAATTTGACCCCTGGTTGGAGAGAGGGCCCAAGTTGAGT 2020  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
Db 2021 GGTGGGCAAGAGCAGAGGATCGCCATTGTCAGCTGCCCTGTTCCGAAACCCCAAGATCCTC 2080  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
Db 2081 CTCGCTGATGAGCCACGTCAGCTTGGACACAGAAGCGACAGCTGGTTGAGTGGCT 2140  
Qy 122 LeuGluLysAlaSerLysGlyArgThrIleValValAlaHisArgLeuSerThrIle 141  
Db 2141 CTCGATAAGCCAGAAAGTCGACACCAATGTCATAGCTCATCGTTGTCACAGTT 2200  
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlnLysGlyAlaHis 161  
Db 2201 CGTAATGCTGACGTCACTCGTGGTTTCGATGATGGAGTCAATTTGTGGAGAAAGGAATCAT 2260

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
Db 2261 GATGAATCATGAAAGAGAAAGGCGATTCTTCAAATTTGTCAACATGCGACACACAGGA 2320  
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrThrGlu 195  
Db 2321 AATGAAGTTGAATTAGAAATGCACCTGATGAATCCAAAGTGAATGATGCTTCTTGGAA 2380  
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
Db 2381 ATGCTCTCAAAATGATTCAAGATCACTTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440  
Qy 210 Lys-----SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluLeuSerLeu 228  
Db 2441 CGTGATCACAACCCCAACAGCAAGAGCTTAGTACAAAGAGGCTCTGGATGAAAGTATA 2500  
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248  
Db 2501 CCTCAGTTTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCTTATTTTGT 2560  
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
Db 2561 GTTGGTGTATTTTGTGCCATTATAAATGAGGCGCTGCAACAGCATTTCGAATAATATT 2620  
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287  
Db 2621 TCAAGATTTATAGGGGTTTTTACAAGAAATTGATGATCTCTGAAACAAACGACAGAACT 2680  
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
Db 2681 AACTTGTTCACATATGTTCTAGCCCTTGGAAATATTCTTTTATTACATTTTTCCT 2740  
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
Db 2741 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAGCGGCTCGGATACATGTT 2800  
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
Db 2801 TTCCGATCCATGCTCAGACAGGATGTGATGTTGTTGTATGATGACCTTAAACACCACTGA 2860  
Qy 348 GlyLeuThrThrIleAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
Db 2861 GCATTGACTTACGAGCTCGCCAAATGATGCTCTCAAGTTAAAGGGGCTATAGTTCCAG 2920  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 2921 CTTCGTGTAATTCACGAGATATAGCAATTCCTGGACAGGAGTAATATATATCCTTCATC 2980  
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2981 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATGTACCCATCATTCGAATAGCAGGA 3040  
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
Db 3041 GTTGTGAAATGAAATGTTGTCTGGCAAGCAGCTGAAAGATAAGAAAGAACTAGAAAGT 3100  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
Db 3101 GCTGGAGAGATCGCTACTGAAACCAATAGAAAATTCCTCGAACCCGTTGTTCTTGTGACTCAG 3160  
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
Db 3161 GAGCAGAAAGTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220  
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 3221 AGGAAAGCACATCTTTGGAATTTACATTTCTTCACCAGGCAATGATGATTTTTC 3280  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 3281 TATGCTGATGTTTCGGTTTGGAGCTTACTTTGGTGGCACAATAAATAATAGCTTTGAG 3340  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527

```
Db 3341 GATGTTCTGTAGTATTTCAGCTGTGTCTTTGGTGGCCATGCGCTGGGCAAGTGAGT 3400
Qy 528 ValLeuAlaProGluThrSerLysAlaLysSerGlyValAlaHisLeuPheAlaLeuLeu 547
Db 3401 TCATTTGCTCCTGACATATGCCAAAGCAAAATATACAGAGCCACATCATCATGATCAT 3460
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3461 GAAAAAACCCCTTTGATTGACAGTACAGCAGGAAGGCCAATGCGGACACATTTGGAA 3520
Qy 568 GlyAsnLeuGluThrArgGluValSerPhePheProCysArgProAspValPheIle 587
Db 3521 GGAATGTACATTTGCTGAAGTTGATTCACTATCCACCCGACCGGACATCCCAAGTC 3580
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyThrValAlaPheValGlySerSer 607
Db 3581 CTTGAGGACTGAGCTGGAGGTGAAGAAGGCCACACGCTGGCTCTGGTGGGACAGT 3640
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuThrAspProValGlnGly 627
Db 3641 GGTCTGGGAGAGCACAGTGGTCCAGCTCTCGAGCGGTTCTACGACCCCTTGGCAGGG 3700
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3701 AAAGTGTCTGTCATGGCAAGAAATAAAGCGACTCAATGTTCAAGTGGCTCCGAGCACAC 3760
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3761 CTGGGATCGTGTCCAGGAGCCCATCTGTTTGGTGCAGCATTTGCTGAGAACATTTGCC 3820
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3821 TATGGACACACAGCGGGTGGTGTACAGGAAGAGATCTGTAGGGCAGCAAGAGGGCC 3880
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysThrAsnThrGlnValGlyLeuLys 707
Db 3881 AACATACATGCTTCATCGAGTCACTGCCTTAATAATATAGCACATAAGTAGGAGACAAA 3940
Qy 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3941 GGAACCTCAGCTCTCTGTGTGCCAGAACACACGCAATTCGCCATGCTGTGCCCTTGTAGA 4000
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4001 CAGCCTCATATTTGCTTTGGTGAAGCCACGTCAGCTCTGATACAGAAAGTGAAG 4060
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4061 GTTGTCCAGAGCCCTGGACAAACCCAGAGAGGGCGCACCTGCATTTGATTGCTCAC 4120
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLys 787
Db 4121 CGCCTGTCCACCATCCAGATGACAGACTTAATAGTGTGTTTCAGAAATGCGCAGTCAAG 4180
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleThrPheLysLeuValAsn 807
Db 4181 CAGATGGCAGCATTCAGCAGTGTCTGGCAGAGAGGCAATCTATTTTCAATGTCAGT 4240
Qy 808 AlaGln 809
Db 4241 GTCCAG 4246
```

## RESULT 12

```
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.823
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/168,621
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,433
; FILING DATE: 27-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Karta, Glenn E.
; REGISTRATION NUMBER: 30,649
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301 527-2058
; TELEFAX: 301 208-6997
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-461-823-1

Alignment Scores:
Pred. No.: 6,35e-257 Length: 2726
Score: 2280.00 Matches: 435
Percent Similarity: 76.55% Conservative: 159
Best Local Similarity: 56.06% Mismatches: 168
Query Match: 55.90% Indels: 14
DB: 1 Gaps: 4

US-09-873-409-2 (1-812) x US-08-461-823-1 (1-2726)
Qy 49 MetGluArgAlaAlaArgGluAlaAlaAlaTyrAspPheIleMetGluPheProAsnLys 68
Db 2 ATTGAGAAAGCTGTCAAGGAAGCCAAATGCTATGACTTTATCATGAACTGCTCATATAA 61
Qy 69 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 89
Db 62 TTTGACACCTGTTGGACAGAGAGGGGCCAGTTGAGTGTGGGCGAAGCAGAGGATC 121
Qy 89 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleuAspGluAlaThrSer 108
Db 122 GCCATTGCGAGCTGCTGTTGCGAACCCCAAGATCTCTCTGCTGGATGAGGCCAGTCA 181
Qy 109 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 128
Db 182 GCCTTGGACACAGAAAGCCAGAGCTGGTTCAGGTGGCTCTGGATAGGCCAGAAAGGT 241
Qy 129 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 148
Db 242 CGGACCAACCATTTGTAGTATCGTTTGTCTACAGTTGCTGAATGCTGACGTCATCGCT 301
Qy 149 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 168
Db 302 GGTTCGATGATGGAGTCATTGTGGAGAAAGGAATCATGATGATGAAAGAGAAA 361
```

Qy	515	AlaIlelaTyRGlyAlaMetAlalleGlyLysThrLeuValLeuAlaProGluTySer	533
Db	1442	GCTGTGTCCTTTGGTGCATGGCCGGGGGCAAGTCAGTTCACTTTGGTCCCTGATATGCC	1501
Qy	535	LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp	554
Db	1502	AAAGCCAAATATACAGAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGAC	1561
Qy	555	SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu	574
Db	1562	AGCTACAGCAGCGAAGGCCTTAATGCCGAAACACATATGGGAAGGAAATGTCACATTTGGTGAA	1621
Qy	575	ValSerPhePheTyRProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer	594
Db	1622	GTGTGTTTCAACTATCCACCCGACCGGACATCCCGAGCTCTTCAGGGACTGAGCCCTGGAG	1681
Qy	595	IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer	614
Db	1682	GTGAAGAAGGCCAGACGCTGGCTCTGTTGGGCAGCAGTGGCTGTGGGAAGACGACACGTG	1741
Qy	615	ValGlnLeuLeuGlnArgLeuTyRAspProValGlnGlyGlnValLeuPheAspGlyVal	634
Db	1742	GTCCAGCTCCTCGAGCGGGTCTACGACCCCTTGGCAGGGGAAAGTCTGCTTGATGGCAAA	1801
Qy	635	AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu	654
Db	1802	GAATAAAGCGACTGAATGTTTCAGTGGCTCCGAGACACCTGGGACATCGTGTCCAGGAG	1861
Qy	655	ProValLeuPheAsnCySserIleAlaGluAsnIleAlaTyRGlyAspAsnSerArgVal	674
Db	1862	CCCATCTGTTTGCATCGACATTGCTGAGAACATTTGCCATATGGAGACAACAGCCGGGTG	1921
Qy	675	ValProLeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu	694
Db	1922	GTGTCA CAGGAAGAGATCGTGGAGGCGAGCAAGAGGCGCAACATACATCGCTTCATCCGAG	1981
Qy	695	GlyLeuProGluLysTyRAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly	714
Db	1982	TCACTGCTTAATAATATAGCACTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGGTGGC	2041
Qy	715	GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu	734
Db	2042	CAGAAACAACGCATTTGCCATAGCTCGTGCCCTTGTGTAGACAGCTCATATTTGCTTTTG	2101
Qy	735	AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisIleLeuAsp	754
Db	2102	GATGAAGCCACGTGCGCTCGGATACAGAAAGTGAAAGGTTGTCCAGAAAGCCCTCGAC	2161
Qy	755	LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn	774
Db	2162	AAAGCCAGAGAGCCGCACCTGCATTTGATTTGTCACCGCTGTCCACCATCCAGAAAT	2221
Qy	775	AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu	794
Db	2222	GCAGACTTAATAGTGTGTTTCAGATGTCAGATGTCAGAGTCAAGGAGCATGGCAGCATCAGCAG	2281
Qy	795	LeuLeuArgAsnArgAspIleTyRLeuLysLeuValAsnAlaGlnSer	810
Db	2282	CTGTGGCACAGAAAGCATCTATTTTTCAATGGTCAGTGTCCAGGCT	2329

RESULT 13

US-08-996-545-1

; Sequence 1, Application US/08996545

; Patent No. 5928898

; GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.

; APPLICANT: de Waard, Maarten A.

; APPLICANT: Peery, Robert B.

; APPLICANT: Andrade, Alan C.

; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of

; TITLE OF INVENTION: Aspergillus nidulans

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company  
 STREET: Lilly Corporate Center  
 CITY: Indianapolis  
 STATE: Indiana  
 COUNTRY: U.S.  
 ZIP: 46285  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/996,545  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Webster, Thomas D.  
 REGISTRATION NUMBER: 39,872  
 REFERENCE/DOCKET NUMBER: X-11766  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 317-276-3334  
 TELEFAX: 317-276-2763  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4002 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..4002  
 US-08-996-545-1

## Alignment Scores:

Pred. No.: 1,87e-187 Length: 4002  
 Score: 1690.50 Matches: 365  
 Percent Similarity: 61.52% Conservative: 152  
 Best Local Similarity: 43.47% Mismatches: 293  
 Query Match: 41.44% Indels: 31  
 DB: 2 Gaps: 7

US-09-873-409-2 (1-812) x US-08-996-545-1 (1-4002)

Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
 Db 1465 TTGCTGATGGCCATGATGATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCT 1524  
 Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnilleLysTyrGly 40  
 Db 1525 TTGCTTAGCAGGAGCGCTGTTCTTTTGGCAGCAGATTTATAGAATAATTAGGACGGT 1584  
 Qy 41 -----ArgAspAspValThrAspGluGluMetGluArg 51  
 Db 1585 CTCATCGGCACAAAGTACGAGATGAATCCGAGGATGAAGTCCGGGAACTCATCGAGAAC 1644  
 Qy 52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71  
 Db 1645 GCGGCAAAATGGCGAATGCTCATGACTTTTATTACTGCCTTGCCTGAAGGTTATGAGACC 1704  
 Qy 72 LeuValGlyGluLysAlaGlnMetSerGlyGlnLysGlnLysGlnAlaIleAla 91  
 Db 1705 AATGTTGGGAGCGGTGGCTTCTCTTTCAGTGGCCAGAACAGCGCATTTGCAATGCC 1764  
 Qy 92 ArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAsp 111  
 Db 1765 CGTGGCGTGTGTAGTACCCCAAAATCTGCTCTGTGATGAGTACTTCGGCTTGGAC 1824  
 Qy 112 SerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThr 131  
 Db 1825 ACAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAATCTACT 1884

Qy 132 IleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLys 151  
 Db 1885 ATTGTGATCGCTCATCGCTTTCCAGCATCAAAACGGCGCAACACATTGTGGTTCGTGTC 1944  
 Qy 152 AspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 171  
 Db 1945 AATGGCAAAATTGCTGAACAAGAACTCAGATGAATTTGTTGACCGCGAGCGCTTAT 2004  
 Qy 172 TyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThr 191  
 Db 2005 CGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAAGCTCAGCGCTTGGAGGAC 2064  
 Qy 192 TyrSerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValLysSerIle 209  
 Db 2065 GCCACGCTGAGGATCTCAGGAATGCAGATATTTGCCAAATCAAAACTGCGTCAAGCGCA 2124  
 Qy 210 LysSerAspPhe-----IleAspLys-----AlaGluGluSerThr 221  
 Db 2125 TCATCGGATCTCGACGGAACCCACACCATTTGACCGCAGCGGACCCACCAATGCTGTT 2184  
 Qy 222 GlnSerLysGluIleSer-----LeuProGluValSerLeuLeuLys 235  
 Db 2185 TCACGCGGATTTCTTTAAAGACCCCGGAAACCAACTCCGAAATACTCATTTATGGACG 2244  
 Qy 236 IleLeuLys-----LeuAsnLysProGluTrpProPheValValLeuGlyThr 251  
 Db 2245 CTGCTCAAAATTTGTTGCTTCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCT 2304  
 Qy 252 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIle 271  
 Db 2305 GTCTTCTCAGTGTAGCTGTGGTGGCCAAACCCACCAAGCAGTGTATATGTAAGCC 2364  
 Qy 272 IleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIle 289  
 Db 2365 ATCAGCACACTCTCGCTCCAGAAATCACAATATAGCAAGCTTCGACATGATCGGATTTTC 2424  
 Qy 290 TyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGly 309  
 Db 2425 TGCTCATTTGATGTTCTTCGTTGTTGTTATCATTCAGTTTATCAGCGCAGTCAACCAATGGT 2484  
 Qy 310 LeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLys 329  
 Db 2485 GCTGCATTTGCCGTATGCTCCGAGACTTTATTCGTCGCGCAGAGACACTGCTTTTCGG 2544  
 Qy 330 AlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeu 349  
 Db 2545 ACATATCTCCGTCAGACATTTGCTTTTTCAGCAAGAGAGAAATAGCCCGGCTCTG 2604  
 Qy 350 ThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369  
 Db 2605 ACCTCTTCTCTCCACCGAGACGAGCATCTCTCCGTTGTAGCGGTGTGACTCTAGGC 2664  
 Qy 370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389  
 Db 2665 ACGATCTTGATGACCTCCACGACCTTAGGAGCGGCTATCATTTATTCCTCGCGATTTGGG 2724  
 Qy 390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409  
 Db 2725 TGAATATGGCTTGTGTTGTTGATCTCGGTTGTCGCGGTTCTCTCGCATCGGTTTCTAC 2784  
 Qy 410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLysHisAlaGly 429  
 Db 2785 CGATTTCTATGCTAGCCCACTTTCAATCAGCTCCAGCTCCAGGTTGTTATGAGGATCTGCA 2844  
 Qy 430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449  
 Db 2845 AACTTTGCTTGCAGGCTACATCGTCTATCCGACAGTTCGCTCATTAACCCGGAAGG 2904  
 Qy 450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469  
 Db 2905 GATGTCGGGAGATTATCCATGCCAGCTTGACGCAAGGAGGACCACTCTAATCTCT 2964



```
QY 470 AlaGlnIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
Db 2965 GTCTTGAGGTATCCCTGTTATATGCGTCGTCGAGCACCTTGTCTTCTCTCGGTGG 3024
QY 490 AlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMet 509
Db 3025 CTCGGGTTTGTGACGAGGACCATCTTGTGTCACACGAGTATGACATTTTCCGCTTC 3084
QY 510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
Db 3085 TTTGTTGTTCTCCGAGATCTCTTGTGTCTCAATCCGCGGCACCGCTCTTTCTTT 3144
QY 530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLys 549
Db 3145 GCACACGATGGCAAGCGAAGATGCGCGCGCGAATTCGACGACTGTTTCGACCGA 3204
QY 550 LysProAsnIleAspSerArgSerGlnGluGlyLysValProAspThrCysGluGlyAsn 569
Db 3205 AGCCCAAAATTGATNACTGCTCTGAAGAGGCGGAGAGCTCGAAACGGTGAAGTGAA 3264
QY 570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
Db 3265 ATCGAATTTAGGAACGTGACATTCAGATACCCGACCGCCAGACAGCTGTCTGCGC 3324
QY 590 GlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
Db 3325 GCGTTGACCTGACCGTGAAGCTGACAAATATGTTGCGCTTGTGCGACCCAGCGTGT 3384
QY 610 GlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyVal 629
Db 3385 GGCAAGATGACCACTTGCATTTGCTGTGAGCGCTTTTACGATGCGATTGCGGTCATC 3444
QY 630 LeuPheAspGlyValAspAlaLysGluLeuAsnValClnTrpLeuArgSerGlnIleAla 649
Db 3445 CTGTGTGAGGGAAGACATAGTAACCTAATATCACTCTACCGAGCTTCTGTGCA 3504
QY 650 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
Db 3505 CTGGTCAGCGAGCGGACACATGTACACGGGCACCATCAAGGAAACATCTTACTTGT 3564
QY 670 AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnIle 689
Db 3565 ATGTGCAAGATGACGTACCGAAGAAATCTTGATTAAAGCTTGAAGGACGCTAATATC 3624
QY 690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
Db 3625 TACGACTTCATATGCTCGCTCCGAGGGCTTTAATACAGTTGTTGGACAGGAGGCG 3684
QY 710 GlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysPro 729
Db 3685 ATGTTGCTGCGGGCCCAAGCAACGTTGTGGCCATTGCGCGAGCCCTTCTTCGGGATCC 3744
QY 730 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
Db 3745 AAAATCTTCTTCTCGATGAAGCGACGTGAGCCCTCGATCCGAGTCAGAAAGGTGCTC 3804
QY 750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
Db 3805 CAGCGGCTTTGATGCTGCGCGAGCGGCGGACCAATCGCCGTTGCACACCGACTC 3864
QY 770 SerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGln 789
Db 3865 AGCAGATTCAAAAGCGGAGGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGC 3924
QY 790 GlyThrHisGlnLeuLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
Db 3925 GGAACGCACAGCGAAGTGTGTCAGAAAAGGCGCGGTACTACGAGTGTGTCACTTTCAG 3984
QY 810 SerVal 811
Db 3985 AGCTTG 3990
RESULT 14
```

```
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:
Pred. No.: 1,87e-187 Length: 4002
Score: 1690.50 Matches: 366
Percent Similarity: 61.52% Conservative: 152
Best Local Similarity: 43.47% Mismatches: 293
Query Match: 41.44% Indels: 31
DB: 2 Gaps: 7

US-09-873-409-2 (1-812) x US-08-996-545-3 (1-4002)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1465 UUGUGGAGGCGCAUGCAUCAAGGACCUCAUCCGCGUGGCUUGCCCAACAGAUUCU 1524
QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
Db 1525 UUGGUUAGCCAGGAGCGUUCUUUUGGCAGCAGCAUUAUUAAGAAUUAUAGGCACGGU 1584
QY 41 -----ArgAspAspValThrAspGluGluMetGluArg 51
Db 1585 CUCAUCCGACAAAGUACGAGAAUCCGAGGAUAAAGUCCGCGGAACUACUACGAGAAC 1644
QY 52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71
Db 1645 GCGGCAAAAUUGCGAUGCUCAUGACUUUAUUAUUGCUUGCCUAGGUAUAGAGACC 1704
QY 72 LeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAla 91
```





```
Qy 350 ThrThrLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369
Db 2605 ACCTCTTCTGTCACCGACGACGAAGCATCTCTCCGGTGTAGCGGTGCTACTAGGC 2664
Qy 370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389
Db 2665 ACGATCTTGATGACCTCCACGACCTTAGAGCGGTATCATATTATGCGCTCGCATGGG 2724
Qy 390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409
Db 2725 TGGAAATGGCGCTTAGTTGATCTCGGTGTGCGGTCTCTCGGCATGCGGTCTTCTAC 2784
Qy 410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGly 429
Db 2785 CGATTCATATGCTAGCCAGCTTCAATCACGCTCCAAAGCTTGCTTATGAGGGATCTGCA 2844
Qy 430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449
Db 2845 AACTTTGCTTCGAGGCTACATCGTCTATCCGCACAGTTGCGTCAATTAACCCGGGAAAG 2904
Qy 450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469
Db 2905 GATGCTGGAGATTTACATGCCAGCTTCACGCACAGCAGCAGCATGCTAATCTCT 2964
Qy 470 AlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
Db 2965 GTCTTGAGGTCACTCCCTGTTATATGCGTCGTCGAGGCACTTGTCTTCTTCGCGTTGCG 3024
Qy 490 AlaGlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMet 509
Db 3025 CTCGGGTTTGTGTCGAGGAGGACACTTCTTGCTCACCACGAGTATGACATTTCCCGCTTC 3084
Qy 510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
Db 3085 TTGTGTTTCTCCGAGATCTCTTTGGTCTCAATCCGCGGGCACCGTCTTTTCTCTTT 3144
Qy 530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLys 549
Db 3145 GCACGACATGGGCAAGCGGAAGATCCGCGCCGCGAATTCGACGACTGTTCCGACCGA 3204
Qy 550 LysProIleIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlyGlyAsn 569
Db 3205 AAGCCACAAATTGATACTGCTCTGAAGAGGCGGAGAGCTCGAAACGGTGGAAAGGTGAA 3264
Qy 570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
Db 3265 ATCGAATTTAGAACGTCACCTCAGATACCCGACCCGCCGACGACGCTGCTCGGC 3324
Qy 590 GlyLeuSerSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
Db 3325 GGCCTTGACCTGACCGTGAAGCTGGACAAATATGTTGCGCTTGTCCGACCCAGCGGTTGT 3384
Qy 610 GlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnVal 629
Db 3385 GGCAGAGATACCCACCATTCGCTTACGCGCTTTTACGATGCGATTCGCGGCTCCATC 3444
Qy 630 LeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla 649
Db 3445 CTGTGTTGATGGGAAGGACATAACTAACTAACTAACTAACTAACTAACTAACTAACTAA 3504
Qy 650 IleValProGlnGluProValPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
Db 3505 CTGCTAGCCAGGAGCCGACACTGTACCAGGGCCACCATCAAGGAAACATCTTACTTGGT 3564
Qy 670 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIle 689
Db 3565 ATTCTCGAAGATGACGTACCCGGAAGAAATTTCTGATTAAAGGCTTCGAAGGACGCTAATATC 3624
Qy 690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
Db 3625 TACACTTCATCATGCTGCTCCCGAGGGCTTTAATACAGTTGTTCGAGCAGGAGGAGGC 3684
Qy 710 GlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysPro 729
```

```
Db 3685 ATGTTGTCTGGCGGCCCAAAAGCAACGTTGTGGCCATTGCCGAGCCCTTCTGGGATCCC 3744
Qy 730 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
Db 3745 AAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTGTC 3804
Qy 750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
Db 3805 CAGCGGCTTTGGATGCGCTGCCGAGGCGGAACCAATCGCCGTTGCACACCGACTC 3864
Qy 770 SerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGluGln 789
Db 3865 AGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGCAAGATCGTCGAAAGC 3924
Qy 790 GlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
Db 3925 GGAACGCACACGCAACTGGTCCAGAAAAGGCGCGGTACTACGAGCTGGTCAACTTCAG 3984
Qy 810 SerVal 811
Db 3985 AGCTTG 3990
```

Search completed: March 31, 2003, 15:23:06  
Job time : 266.913 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 295.562 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-2

Perfect score: 4079

Sequence: 1 MVDENDIRALNVRHYRDHIG.....QELLNRDIYFKLVNAQSVQ 812

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US09873409/runat\_27032003.115424.19290/app\_query.fasta.1.7544  
-DB=Published Applications NA -QMT=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat\_27032003.115424.19290  
-NCPUS=6 -ICPU=3 -NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications NA:
- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:
  - 2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq:
  - 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:
  - 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:
  - 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:
  - 6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:
  - 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:
  - 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:
  - 9: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:
  - 10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:
  - 11: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:
  - 12: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:
  - 13: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:
  - 14: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4079	100.0	2856	10	US-09-873-409-10
2	4079	100.0	3177	10	US-09-873-409-12
3	4079	100.0	3621	10	US-09-873-409-14
4	4079	100.0	3702	10	US-09-873-409-13

5	3456	84.7	2066	10	US-09-873-409-9	Sequence 9, Appli
6	2465.5	60.4	3924	10	US-09-880-107-2299	Sequence 2299, Ap
7	2441	59.8	4369	10	US-09-769-097-1	Sequence 1, Appli
8	2441	59.8	4425	10	US-09-769-097-3	Sequence 3, Appli
9	2438	59.8	4317	9	US-10-044-671-1	Sequence 1, Appli
10	2436	59.7	3860	10	US-09-866-866A-1	Sequence 1, Appli
11	2436	59.7	3860	10	US-09-866-866A-3	Sequence 3, Appli
12	2436	59.7	4643	9	US-10-072-621-2	Sequence 2, Appli
13	2436	59.7	8630	10	US-09-306-417-1	Sequence 1, Appli
14	2436	59.7	8630	10	US-09-306-417-2	Sequence 2, Appli
15	2431.5	59.6	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
16	2425	59.5	4788	10	US-09-866-866A-7	Sequence 7, Appli
17	2404.5	58.9	4189	10	US-09-866-866A-5	Sequence 5, Appli
18	2372	58.2	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
19	2013.5	49.4	5010	10	US-09-917-800A-483	Sequence 3, Appli
20	1694	41.5	4175	10	US-09-749-340-3	Sequence 1, Appli
21	1690.5	41.4	4002	9	US-09-758-828-1	Sequence 3, Appli
22	1690.5	41.4	4002	9	US-09-758-828-3	Sequence 7, Appli
23	1678.5	41.1	3512	10	US-09-749-340-7	Sequence 2237, Ap
24	1669.5	40.9	3861	9	US-09-938-842A-2237	Sequence 5, Appli
25	1617	39.6	2698	10	US-09-749-340-5	Sequence 263, App
26	1550.5	38.0	3861	9	US-09-938-842A-263	Sequence 8, Appli
27	1538.5	37.7	2681	10	US-09-749-340-8	Sequence 2, Appli
28	1452.5	35.6	4653	9	US-10-101-388-2	Sequence 1, Appli
29	1287.5	31.6	8036	9	US-10-101-388-1	Sequence 10, Appli
30	1116.5	27.4	3792	9	US-09-882-694-10	Sequence 9, Appli
31	1032.5	25.3	3999	9	US-09-882-694-9	Sequence 59, Appli
32	866	21.2	7296	10	US-09-070-927A-59	Sequence 12, Appli
33	864	21.2	9047	10	US-09-070-927A-12	Sequence 4, Appli
34	836.5	20.5	1810	10	US-09-749-340-4	Sequence 16, Appli
35	828	20.3	1940	9	US-09-738-626-1251	Sequence 1251, Ap
36	820.5	20.1	3741	9	US-09-873-409-15	Sequence 15, Appli
37	812.5	19.9	2021	10	US-09-873-409-15	Sequence 255, App
38	807	19.8	3601	10	US-09-070-927A-255	Sequence 173, App
39	788	19.3	12438	10	US-09-070-927A-173	Sequence 275, App
40	770	18.9	6415	10	US-09-070-927A-275	Sequence 1, Appli
41	769	18.9	64081	10	US-09-790-988-1	Sequence 18, Appli
42	758.5	18.6	2298	9	US-10-156-239-18	Sequence 18, Appli
43	758.5	18.6	2298	10	US-09-795-693-18	Sequence 16, Appli
44	758.5	18.6	3408	9	US-10-156-239-16	Sequence 16, Appli
45	758.5	18.6	3408	10	US-09-795-693-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-09-873-409-10  
; Sequence 10, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-10

Alignment Scores:  
Pred. No.: 0  
Score: 4079.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 10  
Length: 2856  
Matches: 812  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-10 (1-2856)

Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20
Db	418	ATGCTGGATCAGAATGACATCAGAGCTTTAAATGTGGCCATTATCGAGACCATATTGGA	477
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleGlyTyrGly	40
Db	478	GTGGTTAGTCAAGAGCCGTGTTTGTTCGGGACACCATCAGTAAACAAATATCAAGTATGGA	537
Qy	41	ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp	60
Db	538	CGAGATGATGCTGATGAAGAGATCGAGAGAGCAGCAAGGAGCAAAATGCGTATGAT	597
Qy	61	PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet	80
Db	598	TTTATCATGGAGTTTCTTAATAAATTTAATACATGTTAGGGGAAAAAGAGAGCTCAAAATG	657
Qy	81	SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle	100
Db	658	AGTCGAGGGCAGAAACAGAGATCGCAATTGCTGCTGCCTTAGTTTCGAAACCCCAAGATT	717
Qy	101	LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla	120
Db	718	CTGATTTTAGATGAGGCTACGTCCTGCCCTGGATTCAAGAACCAAGTCAGCTGTTCAAGCT	777
Qy	121	AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr	140
Db	778	GCATCGAGAGAGGCGAGCAAAAGTCGCGACTACATCGTGTAGCACACCGACTTTCTTACT	837
Qy	141	IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla	160
Db	838	ATTCGAAGTCGAGATTGATGTGACCCCTAAAGGATGGAAATGCTGGCGGAGAAAGGAGCA	897
Qy	161	HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle	180
Db	898	CATCCTGAACCTAATGGCAAAACGAGGTCATATTATTCATTGTGATGTCACAGGATATT	957
Qy	181	LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer	200
Db	958	AAAAAAGCTGATGAACAGATGGAGTCAATGCACATATCTTACTGAAAGAAAGACCAACTCA	1017
Qy	201	LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSer	220
Db	1018	CTTCTCTGCACTCTGTGAAGACATCAAGTCAGACTTCATTGACAGGCTGAGAAATCC	1077
Qy	221	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	240
Db	1078	ACCCAACTTAAAGAGATAAGCTTCTCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAAC	1137
Qy	241	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	260
Db	1138	AAGCCTGAAATGGGCTTTGTGGTTCCTGGGGACATTTGGCTTCTGTTTCTAAATGGAACTGT	1197
Qy	261	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys	280
Db	1198	CATCCAGATTTTCCATCATCTCTTCGAAAAATTAACCATGCTTTGGAAAAATAGATATAA	1257
Qy	281	ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle	300
Db	1258	ACCACATTTAAAGCATGATGCAGAAAAATTTATTCATGATATTTCGTCATTTTGGGTGTTATT	1317
Qy	301	CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr	320
Db	1318	TGCTTTGCTCAGTTATTTCATCGAGGAGTATTATTTTACGCAGACAGCGGGAATTTTAAACG	1377
Qy	321	MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp	340
Db	1378	ATGAGATTAGACACATGGGCTTCAAGGCATGTTATATCAGGATATTTCCTGGTTTGAT	1437
Qy	341	GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle	360
Db	1438	GAAGAAGAAACAGCAGAGAGCTTGAACAAATATTAGCCATAGATATGACACAAATTT	1497

Qy	361	GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu	380
Db	1498	CAAGGACCAACAGGTTCCAGGATGGCGCTCTTAAACAACAAATGCAACTAACATGGGACTT	1557
Qy	381	SerValIleIleSerPheIleTyGlyTTPGluMetThrPheLeuIleLeuSerIleAla	400
Db	1558	TCAGTTATCATTTCCCTTTATATATGATGGGAGATGACATCTCGATTTCTGATATTGCT	1617
Qy	401	ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnIys	420
Db	1618	CCAGTACTTCCGCTGCAGGAATGATTTGAAACCGCAGACCAATGACTGGATTTTGCCACACAA	1677
Qy	421	AspIysGlnGluLeuIysHisAlaGlyIlyIleAlaThrGluAlaLeuGluAsnIleArg	440
Db	1678	GATAAGCAAGAACTTAAGCATGCTGAAAGAGATAGCAACTGAAAGCTTTGGAGAAATATCGT	1737
Qy	441	ThrIleValSerLeuThrArgGluIysAlaPheGluGlnMetTyrGluGluMetLeuGln	460
Db	1738	ACTATAGTGTCAATACAGGGGAAAAGCCCTTCGAGCAATGTATGAAGAGATGCTTCAG	1797
Qy	461	ThrGlnHisArgAsnThrSerIysIysAlaGlnIleIleGlySerCysTyrAlaPheSer	480
Db	1798	ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCACG	1857
Qy	481	HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln	500
Db	1858	CATGCCCTTTATATATTTCCTATGACAGAGGTTTCGATTTTGGAGCCCTATTAAATTCAA	1917
Qy	501	AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleTyrGlyAla	520
Db	1918	GCTGCAGCAATGACCCACAGAGGCATGTTCTATGATTTTTTACTTGCAATTTGCATATGGAGCT	1977
Qy	521	MetAlaIleGlyIysThrLeuValLeuAlaProGluTyrSerIysAlaIysSerGlyAla	540
Db	1978	ATGCGCCATCGGAAAAACGCTCGTTTGGCTCTCGTAATATTCGAAAGCCAAATCGGGGGCT	2037
Qy	541	AlaHisLeuPheAlaLeuLeuGluIysIysProAsnIleAspSerArgSerGlnGluGly	560
Db	2038	CGCATCTGTTTGCCTTGTGGAAAGAAACCAATATAGACAGCCGAGTCNAGAAGGG	2097
Qy	561	LysIysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrPro	580
Db	2098	AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAACTCTTCTCTCTATCCCA	2157
Qy	581	CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyIysThr	600
Db	2158	TGTCGCCAGATGTTTTCATCTCCGTCGCTTATCCCTCAGTATTTGAGCGAGGAAGACA	2217
Qy	601	ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg	620
Db	2218	GTAGCATTTGTGGGGAGCAGCGCTGTGGGAAAAAGCACCTCTGTTTCAACTTCGCAGAGA	2277
Qy	621	LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaIysGluLeuAsn	640
Db	2278	CTTTATGACCCCGTCGACAGACAAAGTGCTGTTTGTATGGTGGATGGATGCAAAAGATTTGAAT	2337
Qy	641	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	660
Db	2338	GTACAGTGGCTCGTTCCTCCCAATAGCAATCGTTTCTCTCAGAGCCGTGTGCTCTTCAACTGC	2397
Qy	661	SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle	680
Db	2398	AGCATTCCTGAGAACATCGCCTATGCTGTGACCAACAGCCGCTGGTGCTCCATTAGATGAGATC	2457
Qy	681	LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluIlystyr	700
Db	2458	AAAGAGCCGCAATGAGCAAAATATCCATCTTTTATTGAAGGTTCTCCCTGAGAAATATC	2517
Qy	701	AsnThrGlnValGlyLeuIysGlyAlaGlnLeuSerGlyGlyGlnIysGlnArgLeuAla	720
Db	2518	AACACACAAGTTGCACTGAAAGAGGACACAGCTTCTGGCGGCGCAGAAACAAAGACTAGCT	2577

Qy	721	ILeAlaA	gAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla	740
Db	2578	ATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTATTATTGTTGGATAGAGCCACTTCAGCC		2637
Qy	741	LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaAatgThrGlyArg	760	
Db	2638	CTCGATAATGACAGTCAGAGGTTGGTTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG		2697
Qy	761	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal	780	
Db	2698	ACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCCAGATTGTGATAGTGT		2757
Qy	781	LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp	800	
Db	2758	CTGCACAATGCAAGATAAAGGAACAAGGAACTCATCAAGAGCTCTCTGAGAAATCGAGAC		2817
Qy	801	IleTyrPheLysLeuValAsnAlaGlnSerValGln	812	
Db	2818	ATATATTATTAAGTTAGTGAATGCACAGTCAGTGCAG		2853
RESULT 2				
US-09-873-409-12				
; Sequence 12, Application US/09873409				
; Patent No. US20020037522A1				
; GENERAL INFORMATION:				
; APPLICANT: Sayegh, Mohamed				
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glyco				
; FILE REFERENCE: 81994/268611				
; CURRENT FILING DATE: 2001-06-05				
; NUMBER OF SEQ ID NOS: 19				
; SOFTWARE: PatentIn version 3.0				
; SEQ ID NO 12				
; LENGTH: 3177				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: No. US20020037522A1e				
; LOCATION: (198)..(198)				
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C,				
US-09-873-409-12				
Alignment Scores:				
Pred. No.:		0	Length:	3177
Score:		4079.00	Matches:	812
Percent Similarity:		100.00%	Conservative:	0
Best Local Similarity:		100.00%	Mismatches:	0
Query Match:		100.00%	Indels:	0
DB:		10	Gaps:	0
US-09-873-409-2 (1-812) x US-09-873-409-12 (1-3177)				
Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20	
Db	739	ATGGTGGATCGAATGCATCAGAGCTTTAAATGTGGCATTATCGAGCCATATTGGA	798	
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly	40	
Db	799	GTGGTTTAGTCAAGAGCGCTGTTTGTTCGGGACCACCATCAGTAACTATCAAGTATGGA	858	
Qy	41	ArgAspAspValThrAspGluGluMetGluArgAlaAlaAatgGluAlaAsnAlaTyrAsp	60	
Db	859	CGAGATGATGTGATCGATGAAGAGATGGAGAGACGACGAAGGGAAGCAATCGCTATGAT	918	
Qy	61	PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet	80	
Db	919	TTTATCATGGNGTTTCTTAATAATTTAATACATTGGTAGGGAAAAAGGAGCTCAATG	978	
Qy	81	SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle	100	
Db	979	AGTSGAGGCGAGAAACAGAGATCGCAATGTCTGGTGCCTTAGTTTGGAAACCCCAAGATT	1038	

Qy	101	Leu1LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla	120
Db	1039	CTGATTTTAGATGAGGTACGTCCTCCCTGGATTTCAGAAAGCAAGTCAGCTGTGTCAAGCT	1098
Qy	121	AlaLeuGluLysAlaSerLysGlyVargThrThrIleValValAlaHisAsnLeuSerThr	140
Db	1099	GCACCTGGAAAGCGAGCAAGGTCGGACTCAATTCGCTGGTAGCACACCGACTTTCCTACT	1158
Qy	141	IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla	160
Db	1159	ATTTCGAAGTCAGATTGTATGTGACCTTAAAGGATGAATGCTGCGGAGAAAGGACCA	1218
Qy	161	HisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSerGlnAspIle	180
Db	1219	CATGCTGAACATAATGCGAAACGAGGCTCTATATTATTCACTTGTGATGTCACAGGATATT	1278
Qy	181	LysIleAlaAspGluClnMetGluSerMetThrTySerThrGluLysGlyThrAsnSer	200
Db	1279	AAAAAGACTGATGAACAGATGAGTCAATGACATATTCTACTGAAAGAAAGAACCAACTCA	1338
Qy	201	LeuProLeuHisSerValLysSerLysSerAspPheIleAspLysAlaGluGluSer	220
Db	1339	CTTCCTCTGCACCTCTGTGAAGAGATCAAGTCAGACTTCATTGACAGGCTGAGGAATCC	1398
Qy	221	ThrGlnSerLysGluLeSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	240
Db	1399	ACCAATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAAC	1458
Qy	241	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	260
Db	1459	AAGCCTGAATGGCCTTTTCTGTGTTCTGGGGACATTTGGCTCTGTGTTCTAAATTTGGAACCTGTT	1518
Qy	261	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys	280
Db	1519	CATCCAGATATTTTCCATCATCTTTGCAAAAAATTATPAAACATGTTTGGAAATTAATGATAA	1578
Qy	281	ThrThrLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIle	300
Db	1579	ACCACATTAAAGCATGATGCAGAAATTTATTCATGATATTTCGTCATTTTGGGTGTTATT	1638
Qy	301	CysPheValSerTyrrPheMetGlnGlyLeuPheTyrrGlyArgAlaGlyGluIleLeuThr	320
Db	1639	TGCTTTTGTCAGTTATTTCATGAGGGATATTATTTTACGGCAGACAGCGGCGAAATTTTAAAG	1698
Qy	321	MetArgLeuArgHisLeuAlaPheLysAlaMetIleuTyrrGlnAspIleAlaTrpPheAsp	340
Db	1699	ATGAGATTTAAGACACTTTGGCCTTCAAGGCCATGTATTATATCAGGATATTTCCTGGTTTGAT	1758
Qy	341	GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle	360
Db	1759	GAAGAAGAAAACAGCACAGAGGCTTGACNACAAATTATTAGCCATAGATATAGCACAAATT	1818
Qy	361	GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu	380
Db	1819	CAAGGAGCAACAGTTTCCAGGATTTGGCTTCTTAAACAAAAATGCAACTTAACATGGGACTT	1878
Qy	381	SerValIleIleSerPheIleTyrrGlyTrpGluMetThrPheLeuIleLeuSerIleAla	400
Db	1879	TCAGTTATATCATTTTCCCTTTATATATGATGGGAGATGACATTTCCCTGATTTCTGAGTATTGCT	1938
Qy	401	ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys	420
Db	1939	CCAGTACTTCCCTGTACAGGAATGATTGAACCGCAGCAATGACTGGATTTTGCCNACAAA	1998
Qy	421	AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg	440
Db	1999	GATTAAGCAAGAACTTAAAGCATGCTGGAAAGATAGCAACTGTAAGCTTTTGAGAAATATACGT	2058
Qy	441	ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrrGluGluMetLeuGln	460
Db	2059	ACTATAGTGTCTATTAACAAGGAAAAAGCCTTCGAGCAAAATGTATTGAAGAGATGCTTCAG	2118



Qy 461 ThrGlnHisArgAsnThrSerLysAlaGlnIlelleGlySerCysTyrAlaPheSer 480  
Db 2119 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTATGCAATTCAGC 2178  
Qy 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500  
Db 2179 CATCCCTTTATATATTTGGCTATGCGACAGGGTTTCGATTGGAGGCTATTTAAATTCAA 2238  
Qy 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520  
Db 2239 GCTGACGAATGACCCAGAGGGCATGTTTCATAGTTTTTACTGCATTCGATTCGATTCGAGCT 2298  
Qy 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540  
Db 2299 ATGCCATCGAAACACGCTGTTTGCTCTCTGAATATTCAAAGCCAAATCGGGGCT 2358  
Qy 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560  
Db 2359 GCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGCCGCGAGTCAAGAAAGG 2418  
Qy 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheThrPro 580  
Db 2419 AAAAGCCAGACACATGTGAAGGAAATTAGAGTTTCAGAAAGTCTCTTTCTCTATCCA 2478  
Qy 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600  
Db 2479 TGTGCGCCAGATGTTTTCATCTCGTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 2538  
Qy 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620  
Db 2539 GTAGCATTTGTGGGAGCAGCGGCTGTGGAAAGACATCTCTGTTCACATTCCTCAGAGA 2598  
Qy 621 LeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeuAsn 640  
Db 2599 CTTTATGACCCCGTGAAGGACAGTCTGTTTGATGTTGGATGGCAAGAAATTCGAT 2658  
Qy 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660  
Db 2659 GTACAGTGGCTCCGTTCCCAATAGCAATCGTCTCTCAAGAGCCTGTGCTCTTCAACTGC 2718  
Qy 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680  
Db 2719 AGCATTCGTGAGACATCGCTATGTGTGACACAGCCGCTGTGTGCTTGTGATTCAGATC 2778  
Qy 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700  
Db 2779 AAGAAGCCGCAATGACGCAATATCCATCTTTATTGAAGGTCTCCCTGAGAAATAC 2838  
Qy 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720  
Db 2839 AACACAAAGTTGACTCAAGAGGACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 2898  
Qy 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740  
Db 2899 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2958  
Qy 741 LeuAspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760  
Db 2959 CTCGATATGACGTGAGAGGTGGTTTCAGATGCCCTTTGATAAAGCCAGACGGGAAGG 3018  
Qy 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780  
Db 3019 ACATGCCCTAGTGGTCACTCAGAGCTCTCTGCAATTCAGAACGCGAGATTGTATAGTGGTT 3078  
Qy 781 LeuHisGlnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800  
Db 3079 CTGCACAAATGGAAGATAAAGGAACAAGGAATCTCATCAAGAGCTCTCTGAGAAATCGAGAC 3138  
Qy 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812  
Db 3139 ATATATTTTAAGTTAGTGAATGACAGTCAGTGCAG 3174

RESULT 3

US-09-873-409-14

; Sequence 14, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 14

; LENGTH: 3621

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-873-409-14

Alignment Scores:

Pred. No.: 0 Length: 3621  
Score: 4079.00 Matches: 812  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-14 (1-3621)

Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
Db 1183 ATGGTGGATGAGAATGACATCAGACGCTTTAAATGTGCGCATTCAGAGACCATATTGGA 1242  
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
Db 1243 GTGGTGTAGTCAAGAGCGCTGTTTGTTCGGGACCACCATCAGTAACCAATATCAAGTATGGA 1302  
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
Db 1303 CGAGATGATGCTGACTGATGAAGAGATGGAGAGCAGCAGGAGCAATGCGTATGAT 1362  
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
Db 1363 TTTTATCATGAGTTCCTAATAAATTTAATACATTTGTTAGTGGGAAAAAAGAGCTCAAATG 1422  
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
Db 1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTTGCTGCTGCTTAGTTCCGAAACCCCAAGATT 1482  
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
Db 1483 CTGATTTTAGTGGGCTAGCTCTGCTGCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542  
Qy 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140  
Db 1543 GCACCTGGAGAAGCGCAGCAAAAGGTCGAGCTACAATCGTGGTAGCACACCAGCTTTCTACT 1602  
Qy 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
Db 1603 ATTCGAAGTCGATTTGATTGTGACCTTAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1662  
Qy 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
Db 1663 CATGCTGAATTAATGCAAAACGAGGCTCTATATTTACTTGTGTGTACAGGATATT 1722  
Qy 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
Db 1723 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAAAGAAAGCAACTCA 1782  
Qy 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220  
Db 1783 CTTCCCTCTGCATCTCTGTGAAGAGCATCAAGTCAGACTTTCATTGACAGGCTGAGGAATCC 1842

```
QY 221 ThrGlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysLeuLysLeuAsn 240
DB 1843 ACCCAATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATAAAATTTTAAAGTTAAAC 1902
QY 241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
DB 1903 AAGCCTGAATGGCTTTTGTGGTCTGGGACATTTGGCTTCTGTTCTTAATGGAATGTT 1962
QY 261 HisProValPheSerIlelePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
DB 1963 CATCCAGTATTTCCATCATCTTTTGCAGAAATTTAATCAATGTTTGGAAATATGATAAA 2022
QY 281 ThrThrLeuLysHisAspAlaGluIleThrSerMetIlePheValIleLeuGlyValIle 300
DB 2023 ACCACATTAAGCATGATGAGAAATTTATTCATGATATTCGTATTTGGGTGTTAT 2082
QY 301 CysPheValSerThrPheMetGlnGlyLeuPheThrGlyValArgAlaGlyGluIleLeuThr 320
DB 2083 TGCCTTTGTGAGTTATTTTCATGAGGAGTATTTTACGGCAGACAGCGGAAATTTTAAACG 2142
QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuThrGlnAspIleAlaTrpPheAsp 340
DB 2143 ATGAGATTAAAGACATTTGGCTTCAAGCCATGTTATATCAGGATATTCCTGCTGTTGAT 2202
QY 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
DB 2203 GAAAGGAAACACGACAGGAGGCTTGACAAATATTTAGCCATAGATATAGCACAAAT 2262
QY 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
DB 2263 CAAGGAGCAACAGTTCCAGGATTTGGCGTCTTAACACAAATCCAACTAACATGGGACT 2322
QY 381 SerValIleIleSerPheIleThrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
DB 2323 TCAGTTATCATTTCCTTTATATATGATGGAGATGACATTCCTGATTCGTAGATTGCT 2382
QY 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
DB 2383 CCAGTACTTGGCTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGGCAACAAA 2442
QY 421 AspLysGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
DB 2443 GATAAGCAAGAACTTAAGCATCTGGAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2502
QY 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetThrGluGluMetLeuGln 460
DB 2503 ACTATAGTGCATTAACAGGGAAGAAAGCTTCGACCAATGATGAGAGATGCTTCAG 2562
QY 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysThrAlaPheSer 480
DB 2563 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGAAGCTGTTATGCATTCAGC 2622
QY 481 HisAlaPheIleThrPheAlaThrAlaAlaGlyPheArgPheGlyAlaThrLeuIleGln 500
DB 2623 CATGCGCTTTATATTTTGCCTATGACGAGGTTTCGATTTGGAGCTATTTAAATCAA 2682
QY 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaThrGlyAla 520
DB 2683 GCTGGACGAATGACCCAGAGGGGATGTTCTATAGTTTTTACTGCAATTCATATGAGCT 2742
QY 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluThrSerLysAlaLysSerGlyAla 540
DB 2743 ATGCCCATCGAAAAGCTCGTTTGGCTCCTGAAATATTTCCAAAGCAATTCGGGGCT 2802
QY 541 AlaHisLeuPheAlaLeuLeuGluLysLysPheProAsnIleAspSerArgSerGlnGluGly 560
DB 2803 GCGCATCTGTTGCTTGTGTTGGAAAAGAAACCAATATATAGACAGCCGACGCAAGAAGGG 2862
QY 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheThrPro 580
DB 2863 AAAAGCCACACATGTGAAGGAAATTTAGAGTTTCGAGAAGTCTCTTCTCTATCCA 2922
QY 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
```

```
DB 2923 TGTGCCCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGAAAGACA 2982
QY 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
DB 2983 GTAGCATTTGTGGGAGCAGCGGCTGTGGAAAGACACTTCTGTTCACTTCTGCAGAGA 3042
QY 621 LeuThrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
DB 3043 CTTTATGACCCCGTCAGGACAAAGTGTGTTGTATGGTGTGATGTCGCAAAAGAAATTTGAAT 3102
QY 641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
DB 3103 GTACAGTGGCTCCGTTCCTCCAAATAGCAATCTCTCTCAAGAGCTGTGCTCTTCAACTGC 3162
QY 661 SerIleAlaGluAsnIleAlaThrGlyAspAsnSerArgValValProLeuAspGluIle 680
DB 3163 AGCATTTGCTGAGAACATCGCCTATGGTGACACACAGCGGTGTGGTCCATTAGATGAGATC 3222
QY 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysThr 700
DB 3223 AAAGAAGCCGCAATGCGACAAATATCCATCTCTTTTATTGAAGGTCTCCCTCGAGAAATAC 3282
QY 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
DB 3283 AACACACAAGTTGGACTGAAGGAGCACAGCTTCTGGCGGCCAGAAACAAAGACTAGCT 3342
QY 721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
DB 3343 ATGCAAGGCTCTTCTCCAAAACCCAAAATTTTATTGTTGATGAGGCCACTTCAGCC 3402
QY 741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
DB 3403 CTCGATAATGACAGTGAGAAAGTGGTTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAG 3462
QY 761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleValVal 780
DB 3463 ACATGCTTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTGTAGTGGTT 3522
QY 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAsp 800
DB 3523 CTGCACAATGGAAAGATAAAGGACAAAGGAACTCATCAAGAGCTCTCTGAGAAATCGAGAC 3582
QY 801 IleThrPheLysLeuValAsnAlaGlnSerValGln 812
DB 3583 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

## RESULT 4

```
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homolog on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873, 409
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13
```

Alignment Scores:  
Pred. No.: 0 Length: 3702

Score: 4079.00 Matches: 812  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-13 (1-3702)

QY 1 MetValAspGluAsnAspIleAtrGAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
DB 1264 ATGGTGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCAATATCGAGACCAATATTGGA 1323  
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
DB 1324 GTGGTTAGTCAAGAGCCGTGTTTGTTCGGGACCACCATCATGTAACAATAATCAAGTATGGA 1383  
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
DB 1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGACGACAAAGGAAGCAAAATGCGTATGAT 1443  
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
DB 1444 TTTATCATGAGATTTCTCTAATAATTTAATATCATTTGGTAGGGGAAAAGAGGCTCAATG 1503  
QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
DB 1504 AGTGGAGGCGAGAAACAGAGGATCGCAATTTGCTGCTGCTTAGTTTGGAAACCCCAAGATT 1563  
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
DB 1564 CTGATTTTAGTAGGCTACGCTGCTCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623  
QY 121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140  
DB 1624 GCATGGAGAGGGAGCAAGAGTCGACATACAAATCTGTGTAGCACACCCACCTTTCTACT 1683  
QY 141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160  
DB 1684 ATTCGAAGTCGAGATTTGATTGTGACCTTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1743  
QY 161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180  
DB 1744 CATGCTGAACCTAATGGCAAAACGAGGCTATATTAATTCATTTGATGTGCACAGGATATT 1803  
QY 181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200  
DB 1804 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAAGAAAGACCAACTCA 1863  
QY 201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220  
DB 1864 CTTCTCTGACACTCTGTGGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1923  
QY 221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240  
DB 1924 ACCCAATCTAAAGAGATAAGTCTTCTCTGAAAGTCTCTCTATTAATAAAATTTTAAAGTTAAAC 1983  
QY 241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260  
DB 1984 AAGCCTGAATGGCCCTTTGTGGTGTCTGGGACATGTGCTCTGTCTTAAATGGAACTGTT 2043  
QY 261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280  
DB 2044 CATCCAGTATTTTCCATCATCTTTGGCAAAAATATATACCATGTTTGGAAATAATGATAAA 2103  
QY 281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300  
DB 2104 ACCACATTAAGACATGATGCAGAAAATTTATTTCCATGATATTTCGTCAITTTTGGGTGTTATT 2163  
QY 301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320  
DB 2164 TGCCTTGTGCTAGTTATTTTCATGCGGGATTTATTTTACGCGAGAGAGGGGAAATTTTAACG 2223  
QY 321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340

DB 2224 ATCAGATTAAAGCACCTGTCCTTCAAAGCCATGTTATATCAGGATATTCCTCGGTGTGAT 2283  
QY 341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360  
DB 2284 GAAAGGAAACAGACAGAGGCTTTGACAACAATATTAGCCATAGATATAGCACAAATT 2343  
QY 361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380  
DB 2344 CAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAATGCAACTAATCAGGACTT 2403  
QY 381 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 400  
DB 2404 TCAGTTATCATTTCTTTATATATGATGGAGATGACATTCCTGATTCTGAGTATTGCT 2463  
QY 401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420  
DB 2464 CAGATCTTCCGTGACAGAAATGATTGAACCGCAGCAATGACTGGATTTGCCAACAAA 2523  
QY 421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440  
DB 2524 GATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAAAGCTTTGGAGAATAACGT 2583  
QY 441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460  
DB 2584 ACTATAGTGTCTATTAAACAAGGGAAAAGCCTTCGAGCAAAATGTATGAAGAGATGCTTCAG 2643  
QY 461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480  
DB 2644 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATTTGGAAAGCTGTTATGCAATTCAGC 2703  
QY 481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500  
DB 2704 CATGCCCTTTATATATTTTGCCTATGCGAGCGGCTTTTCGATTTGGAGCCTATTATTAATCAA 2763  
QY 501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520  
DB 2764 GCTGGAGCAATCACCCAGAGGCGCATGTTTCATAGTTTTTACTGCAATTCATATGAGCT 2823  
QY 521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540  
DB 2824 ATGGCCATCGAAAAACGCTCGTTTTGGCTCTCTGAATATTTCCAAAGCCAAATTCGGGGCT 2883  
QY 541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560  
DB 2884 GCGCATCTGTTTGGCTTTGTTGGAAGAAACCAAAATATAGACAGCCGCGAGTCAAGAAGGG 2943  
QY 561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580  
DB 2944 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTTCGAGAAGTCTCTTTCTTCTATCCA 3003  
QY 581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600  
DB 3004 TGTGCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAGACA 3063  
QY 601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620  
DB 3064 GTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAAGCACCTTCGTTCAACTTCCTTCAGAGA 3123  
QY 621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640  
DB 3124 CTTTATGACCCCGTCAAGGACAAAGTCTGTTTGTGTTGATGTTGATGATGCAAAAGAAATGAAT 3183  
QY 641 ValGluTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660  
DB 3184 GTACAGTGGCTCCGTTCCCAAAATAGCAATCGTTCTCCAGAGCCTGTGCTCTTCAACTGC 3243  
QY 661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680  
DB 3244 AGCATTTGCTGAGAACATCGCCTATGTGTGACACAGCCGCTGTGCTGCTTAGATGAGATC 3303  
QY 681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700

```
Db 3304 AAGAAGCCCAAAATGAGCAAAATATCCATCTCTTTATTGAAGGTCTCCCTCGAGAAATAC 3363
QY 701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
Db 3364 AACACCAAGTTCGACTGAAGGAGGACACAGCTTCTGGCGGCCAGAAAACAAAGACTAGCT 3423
QY 721 ILAAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
Db 3424 ATTGCAAGGCTCTCTCTCAAAAACCCAAATTTATTGTTGGATGAGGCCACTTCAGCC 3483
QY 741 LeuAspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
Db 3484 CTCGTAATACACAGTGAAGAGGGTTCAGCATGCCCTTGTAAGCCAGGACGGGAAGG 3543
QY 761 ThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
Db 3544 ACATGCTAGTGTCTACACAGGCTCTCTGCAATTCAGAACCGACAGATTGTAGTGGTT 3603
QY 781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
Db 3604 CTGCACAAATGGAAGATAAAGGAACCAAGGAACCTCATCAGAGCTCTCGAGAAATCGAGAC 3663
QY 801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
Db 3664 ATATATTTTAAAGTTAGTGAATGCACAGTCAGTCAG 3699

RESULT 5
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-9

Alignment Scores:
Pred. No.: 0 Length: 2066
Score: 3456.00 Matches: 687
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.73% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-2 (1-812) x US-09-873-409-9 (1-2066)
QY 126 SerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAsp 145
Db 3 AGCAAGGTCCGACTCAATCGTGTAGCACACCGACTTCTACTATTCGAAGTGCAGAT 62
QY 146 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 165
Db 63 TTGATTGTGACCTAAAGGATGGAATCTCGCGGAGAAAGAGACACATGCTGAACTAATG 122
QY 166 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysAlaAspGlu 185
Db 123 GCAAAAACGAGGCTATATATTACTTGTGATGTACAGGATATATAAAAGCTGATGAA 182
QY 186 GlnMetClnSerMetThrTyrSerThrClnArgLysThrAsnSerLeuProLeuHisSer 205
Db 183 CAGATGGAGTCAATGACATATTTACTGAAAGAAAGAACCAACTCCTCTCTGCACTCT 242
QY 206 ValLysSerIleLysSerAspPheIleAspLysAlaGluSerThrGlnSerLysGlu 225
```

```
Db 243 GTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAAATCCACCAATCTAAAGAG 302
QY 226 IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrPro 245
Db 303 ATAAGTCTTCCTGAAGTCTCTCTATTAAAATTTTAAAGTTAAACAAGCTGATGGCCT 362
QY 246 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 265
Db 363 TTTGTGGTTCTGGGACATTGGCTTCTGTTCTAAATGGAACGTGTTTCATCCAGTATTTTC 422
QY 266 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 285
Db 423 ATCATCTTTTGCAAAATATATACCATGTTTGGAAATTAATGATAAAACCATTTAAAGCAT 482
QY 286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db 483 GATGCAAGAAATTTATTCATGATATTCGTCATTTGGGTGTTATTGCTTTGTCAAGTAT 542
QY 306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db 543 TTCATGCAAGGATTTATTTACGCGAGCAGCGGGAATTTTAACGATGAGATTAAGACAC 602
QY 326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer 345
Db 603 TTGGCCTTCAAAAGCCATGTTATATCAGGATATTCGCTGCTGTTTGAATGAAAGGAAACAG 662
QY 346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db 663 ACAGGAGGCTTGACAAATATTTAGCCATAGATATAGCACAAATTCAGAGAGACAGGT 722
QY 366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db 723 TCAGAGTTGGCGTCTTAACACAAATGCACTAAACATGGGACTTTTCAGTTATCATTTCC 782
QY 386 PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db 783 TTTATATATGATGGAGATGACATTCCTGATTTCTGATATTGCTCCAGTACTTGCCTG 842
QY 406 ThrGlyMetIleGluThrAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db 843 ACAGGAATGATTGAAACCCGACGCAATGACTGGATTTGCCAACAAAGATAGCAAGAACCT 902
QY 426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db 903 AAGCATGCTGGAAGATAGCAACTGAAGCTTTGGAGAAATATACGTACTATAGTGTCTATTA 962
QY 446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db 963 ACAAGGGAAAAAGCCTTCGAGCAAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAAT 1022
QY 466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db 1023 ACCTCGAAGAAAGCAGAGATTTATGGAAGCTGTTATGCAATTCAGCATTCGCTTTATATAT 1082
QY 486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db 1083 TTTGGCTATGACAGCGGTTTCGATTTGGAGCTATTTTAATTCAGCTGGACCAATGACC 1142
QY 506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db 1143 CCAGAGGCGCATGTTTCATAGTTTTTACTGCAATTCATATGGAGCTATGGCCATCGGAAAA 1202
QY 526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db 1203 ACCTCGTGTGGCTCTCGAATATTCAAAGCCAAATCGGGGGCTCGCATCTGTTTGGC 1262
QY 546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db 1263 TTGTTGGAAGAAACCAATATATAGACACCGCAGTCAAGAGGGAAGAAAGCCAGACACA 1322
QY 566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
```

```
Db 1323 TGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTCTATCCATGTCGCCCGAGATGTT 1382
Qy 586 PheIleuArgGlyLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db 1383 TTCAATCTCGTGGCTTATCCCTCAGTATTGAGCGAGAAAGACAGTAGCATTTGTGGGG 1442
Qy 606 SerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db 1443 AGCAGCGCTGTGGAAAGACACTTCTGTTCACCTTCGACAGACTTTATGACCCCGTG 1502
Qy 626 GlnGlyGlnValLeuPheAspGlyValAsnAlaLysGluLeuValGlnTyrLeuArg 645
Db 1503 CAAGGCAAGTGTCTTTGTATGGTGTGGATGCAAAAGAAATGTAATGTACAGTGGCTCCGT 1562
Qy 646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db 1563 TCCCAATAGCATCGTTCTCTCAGAGCCTGTGCTCTTCACTGCAGCATTTGCTGAGAAC 1622
Qy 666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAsn 685
Db 1623 ATCGCCTATGGTGACAACAGCGGTGTGGTGCCATTAGATGAGATCAAAAGACCGCAAT 1682
Qy 686 AlaAlaSerIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db 1683 GCAGCAAAATATCCATCTTTTATTGAAGGTCTCCCTGAGAAATACAAACACAGTTGGA 1742
Qy 706 LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db 1743 CTGAAAGAGCAGACGCTTCTGGCGGCGCAGAAACAGACTAGCTATTGCAAGGGCTCTT 1802
Qy 726 LeuGlnLysProTyrIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db 1803 CTCCAAAACCCAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT 1862
Qy 746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db 1863 GAGAAGTGGTTTCAGCATGCCCTTGATAAGCCAGGACGGGAAGACATGCCTAGTGGTC 1922
Qy 766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db 1923 ACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATGTTCTGCAATGGAAG 1982
Qy 786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Db 1983 ATAAAGGAACAAGGAACATCATCAAGAGCTCTCTGAGAAATCGACACATATATTTAAGTTA 2042
Qy 806 ValAsnAlaGlnSerValGln 812
Db 2043 GTGAATGCACAGTCAGTGCAG 2063
```

RESULT 6

US-09-880-107-2299

```
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
; US-09-880-107-2299

Alignment Scores:
Pred. No.: 5,86e-233 Length: 3924
Score: 2465.50 Matches: 478
Percent Similarity: 76.46% Conservative: 149
Best Local Similarity: 58.29% Mismatches: 182
Query Match: 60.44% Indels: 11
DB: 10 Gaps: 3
```

US-09-873-409-2 (1-812) x US-09-880-107-2299 (1-3924)

```
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1395 ATTGATGGCCAGGATATTAGGAACCTTTAATGTAAACATCTCTGAGGGAATCAATTGGTGTG 1454
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1455 GTGAGTCAGGAGCGGTGCTGTTTCCAGCACAAATTTGCTGAAAATATTTTATTATGCCCGT 1514
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1515 GGAATGTAACTATGATGAGATAAAGAGCTGTCAAAGAGGCCAACCCCTATGAGTTT 1574
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1575 ATCATGAATATTACCACAGAAATTTGACACCTCGTGGTGGAGAGAGAGGGCCAGCTGAGT 1634
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1635 GGTGGGCAGAAAGCAGAGGATCGCCATTGACGCTGCCCTCGTTCGCAACCCCAAGATCCTT 1694
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1695 CTGCTGGATGAGCCACGTCAGCAITGGACACAGAAAGTGAAGCTGAGTACAGGAGCT 1754
Qy 122 LeuGlyLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1755 CTGGATAAGCCAGAGAGCGCGGACCACTTGTATAGCACACCGCACTGTCTACGGTC 1814
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlnLysGlyAlaHis 161
Db 1815 CGAATATGCAGATGTCATCGCTGGGTTTGAAGATGGAGTAATTTGTGGAGCAAGAACCCAC 1874
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 175
Db 1875 AGCGAACTGATGAAGAGGAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934
Qy 176 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
Db 1935 AGCCAGATCCAGTCAGAGAAATTTGAACATAATGATGAAAGGCTGCCACTAGATGGCC 1994
Qy 193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys----Ser 211
Db 1995 CCAATGGCTGGAAATCTCGCCTATTATTAGGCATTTACTCAGAAAAACCTTAAAAAATCA 2054
Qy 212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db 2055 CAAATGTGTGAGAGAGCGCTTGTATGTGGAACCGATGGAGCTTGAAGCAAAATGTCACCA 2114
Qy 231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGly 250
Db 2115 GTGCTCTTCTGAAGGCTCTGAACTGAATAAACAGATGGCCCTACTTTGTCGTGGGA 2174
Qy 251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db 2175 ACAGTATGTGCCATTGCCAATGGGGGCTTCAGCCGCAATTTTCAGTCATATTTCTCAGAG 2234
Qy 271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db 2235 ATCATAGCGGATTTTGGACCAGCGCGATGATGATGCGTGAAGCAGCAGAGAGTGCACATATTC 2294
```

```
QY 291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2295 TCTTTGATTTCCTATTCTGGAATATTCTTTTACTTCTCTTCCCTCAGGCGTTTC 2354
QY 311 PheTyrGlyArgAlaGlyGluIleLeuMetArgLeuArgHisLeuAlaPheLysAla 330
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2355 ACCTTTGGGAAGCTGGCAGATCTCCACAGAAAGACTGCGGTCAATGGCTTTTAAAGCA 2414
QY 331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2415 ATGCTAAGCAGACATGAGCTGTTTGATGACCATAAATAACAGTACTGGTGACATTTCT 2474
QY 351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2475 ACAAGACTTGCCACAGATGTCGCCAAGTCCAAAGGAGCCACAGAAACCAGGTGGCGTTTA 2534
QY 371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594
QY 391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2595 CAGTTAACCTATTGCTATTAGCAGTGTTCATATTATTGCTGTGTGTCAGGAATGTGAA 2654
QY 411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlnLeuLysHisAlaGlyLys 430
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2655 ATGAATTTGTTGGCTGGAATGCAAAAGAGATAAATAAGAACTGGAAGCTGCGAAAG 2714
QY 431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2715 ATTGCAACAGAGCAATAGAAATATATTAGACAGATTGTGTCTTTGACCCAGGAAGAAA 2774
QY 451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2775 TTTGAATCATGATGATGTTGAAAAATTGTATGGACCTTACAGGAATCTGTGCAGAGGCCA 2834
QY 471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2835 CACATCTATGGAATTACTTTTAGTATCTCACAAAGCATTTATGTATTCTTCTATGCGGT 2894
QY 491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2895 TGTTTTCGATTTGGTGCATATCTCATTTGTGAATGGACATATGCGCTTCAGAGATGTTAT 2954
QY 511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2955 CTGGTGTCTTCTGCAATTGATTGTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTGCT 3014
QY 531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 550
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3015 CCAGACTATGCTAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAGACAA 3074
QY 551 ProAsnIleAspSerArgSerGlnGluGlyLysPheProAspThrCysGluGlyAsnLeu 570
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAATATA 3134
QY 571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3135 ACATTTAATGAATGCTGTGTTCAACTATCCACCAGCAAAACGTGCCAGTTCAGGGG 3194
QY 591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3195 CTGAGCTGGAGGTGAAGAAGCCAGACACTAGCCCTGTTGGGAGCAGTGGCTGTGGG 3254
QY 611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3255 AAGAGCAGCGTGGTCCAGCTCCTGGAGCGGTCTTACGACCCCTTGGCGGGACAGTGCCT 3314
QY 631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3315 CTCGATGGTCAAGAGCAAGAAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374
```

```
QY 651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3375 GTGTCTCAGGAGCCTATCTATTTCAGCTGCAGCATTTGCCGAGATATTTCCTTATGGAGAC 3434
QY 671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3435 AACAGCCGGTGTATCATCAGGATGAAATTGTAGTGCAGCCAAAGCTGCCAACATACAT 3494
QY 691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3495 CCTTTTCATCGACAGCTTACCCCAAAATATGAAACAAGAGTGGAGATAAGGGAGCTCAG 3554
QY 711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3555 CTCTCAGGAGGTCAAAAACAGAGGATTTGCTATTGGCCGAGCCCTCATCAGACAACCTCAA 3614
QY 731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3615 ATCCCTCTGTTGGTGAAGCTACATCAGCTCTGGATCTGAAAGTGAAGAGTGAAGTTGTCCAA 3674
QY 751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3675 GAAGCCCTGCAGAAAGCCAGAGAGCGCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734
QY 771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3735 ACCATCCAGAATGCAGACTTAATAGTGTGTTTTCAGAAATGGGAGAGTCAAGGAGCATGCG 3794
QY 791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3795 AGCATCAGCAGCTGCTGCAGAAAGGCATCTATTTTTCATATGTCAGTGTCCAGGCT 3854

RESULT 7
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.: 1,79e-230 Length: 4369
Score: 2441.00 Matches: 470
Percent Similarity: 77.04% Conservative: 164
Best Local Similarity: 57.11% Mismatches: 175
Query Match: 59.84% Indels: 14
DB: 10 Gaps: 5

US-09-873-409-2 (1-812) x US-09-769-097-1 (1-4369)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```







```
QY 728 LysProlylLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
DB 3904 CAGCCTCACATCTTACTTCTGGATGAAGCAGACATCAGCTCTGGATACGGAGTGAAG 3963
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
DB 3964 GTGCTTCAGGAAGCGTGGCAAGCCAGGAGGAGCGCCACCTGCAATGTGTGTCGGCAC 4023
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
DB 4024 CGCCTGTCCACCATCCAGAACGCGAGACTTGATCGTGTGTGATTCCAGAACGCGCCAG 4083
QY 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
DB 4084 GAGCAGCGGACCCAGCAGCTGCTGGCCAGAAAGGCATCTATTCTCGATGGTCAGT 4143
QY 808 AlaGlnSer 810
DB 4144 GTCCAGGCT 4152
RESULT 8
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3
Alignment Scores:
Pred. No.: 1,82e-230 Length: 4425
Score: 2441.00 Matches: 470
Percent Similarity: 77.04% Conservatve: 164
Best Local Similarity: 57.11% Mismatches: 175
Query Match: 59.84% Indels: 14
DB: 10 Gaps: 5
US-09-873-409-2 (1-812) x US-09-769-097-3 (1-4425)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1684 ATCCAGCGGACGAGCATCAGGACCATCAATGTGAGGTATCTCGCGGAAATCAATGGGGTG 1743
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1744 GTGAGTCAGAACCCCGTGTGTTGGCCACCAATTCGCCGAAACATTCGCTATGGCCGA 1803
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1804 GAAACGTCACCATGATGATGATAGAGAAAGCTGTCTCAGGAGACCATGCTATGATTTC 1863
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1944 GTGGTTGAAATGAAATGTTGTCTGGCAAGCGCTGAAAGATAAGAAAGAACTAGAAAGT 3003
```

```
DB 1864 ATCATGAACCTCCCCACAAAATTTTGACACCCTGGTGGTGAGAGAGGGCGGAGCTGAGT 1923
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1924 GGGGACACAAAACAGAGGATGCCATTGCGCGGCCCTGTCGCAACCCCAAGATCCTT 1983
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1984 TTGTTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCAAGCCGCTGTTTCAGGCCGT 2043
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIle 141
DB 2044 CTGGATAGGCTAGAGAAGCGCGACCATTTGTGATAGTACCGCTTGTCTACAGTT 2103
QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
DB 2104 CGCAATGCTGAYGTCATTGCTGCTGTTTGTGTTGTGTTGTCATTGTTGGAGCAAGAAATCAT 2163
QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
DB 2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCTATGACTCAGACAGCAGGA 2223
QY 179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----MetThr 191
DB 2224 AATGAAATGGAATTAGGAATGAAGCTTGTGAATCTAAAGAYGGAATGATATGATGAGAC 2283
QY 192 TyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIle 209
DB 2284 ATGCTCTCAAAAGATTCRGATCCAGTCTTAATAAGAGAGAGATCAACTCGCAAAAGCATC 2343
QY 210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluLysSerLeu 228
DB 2344 CGTGGGCCACATGATCAAGACGCGGAACTTAGCACCAAGAGGCTCTGGATCAGCAGTA 2403
QY 229 ProGluValSerLeuLysLysIleLeuLysLysLeuAsnLysProGluTrpProPheVal 248
DB 2404 CTTCCAGCTTCTCTTTTGGCGGATCTGAAAGTTGAAATCAACTGAATGGGCTTATTTTGTG 2463
QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
DB 2464 GTTGGTGTATTTTGTGCCATAATAATGAGGCTTGCAACACGACATTTCTCCATAATTT 2523
QY 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrLeu---LysHisAspAla 287
DB 2524 TCAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAAAATCCAGCGGAGAACAGC 2583
QY 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
DB 2584 AACTGTTTCTTTTATTTGTTCTGATCTCTGGGATCATCTTTTCAATTCGTTTTTCTTT 2643
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
DB 2644 CAGGCTTCACATTTGGCAAAAGCTGGAGAGATCTCTCACAAGGACTCCGATACATGGTTC 2703
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
DB 2704 TTCAAATCCATGCTGAGACAGACATAAGCTGTTTGTATGACCCCTAAATAAACACACAGCA 2763
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
DB 2764 GCGCTGACACACCGGCTTGCCAAATGACGCTGCTCAAGTGAAGGAGGCTACAGGGTCTAG 2823
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
DB 2824 CTTCGTTTATTAACCAAGACATAGCAAAATCTTGGGACAGGCTCATCATATCCCTGATC 2883
QY 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
DB 2884 TACGGCTGCAATTTGACACTTTTACTCTTAGCAATTTGTTCCCATCATCTGCTATAGCAGGA 2943
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
DB 2944 GTGGTTGAAATGAAATGTTGTCTGGCAAGCGCTGAAAGATAAGAAAGAACTAGAAAGT 3003
```

```
QY 428 AlaGlyIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3004 TCTGGGAAGATCCGTACAGAAAGCAATGAAATTTGCGACTGCTCTTTGACTCGG 3063
QY 448 GluIysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3064 GAGCAGAAAGTTTGAAACTATGTATGCCAGAGCTTGCAGATACCACATACAGAAATGCTTGG 3123
QY 468 LysIysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 3124 AAGAAAGCGCACGTCTTGGGATCACCTTTCTCTTCCACCGGCCATGATGATTTCTCC 3183
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 3184 TATGCTCTTGTTCGGTGTGATGCTTCTTCTTCCACCGGCCATGATGATTTCTCC 3243
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyIysThrLeu 527
Db 3244 AATGTTCTGTAGTATTTCTCAGCTATTGTTCTTGTGCGCATGGCAGTGGCGCAGGTCAGT 3303
QY 528 ValLeuAlaProGluTyrSerLysAlaIysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 3304 TCATTCGCTCTGACTACGGGAAAGCCAAAGTCTCGGCATCCACATCATCAGGATCAT 3363
QY 548 GluIysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3364 GAGAAATCTCTGAGTATGACGTACAGACAGGAGGCTTGAAGCCATATATGTTGGAA 3423
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3424 GGAATGTGAAATTTAATGGAGTCATGTTCAACTATCCACCCGACCAACATCCAGTG 3483
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3484 CTTCAGGGCTGAGCCTAGAGTGAAGAAAGGCGACAGCTGGGCCCTCGTGGCGCAGCAGT 3543
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3544 GCGTCGGGAAGATGACAGTGTCCAGTCTGCTTGGGCTTCTATGACCCCATGCGCGGA 3603
QY 628 GlnValLeuPheAspGlyValAlaPalaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3604 ACAGTGTCTTAGATGGCAAGAAATAAAGCAACTCAATGTCAGTGGCTCCGGGCCAC 3663
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3664 CTGGGCATTTGTGTCCAGAGCCCATCTGTTGACTGCAGCATCGCGGAGAACATTGCC 3723
QY 668 TyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3724 TACGGAGACAAACAGCGGTGTCGTCTCATAGGAGATCGTGAAGGAGCCAGCAGAGGCC 3783
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3784 AACATCCACAGTTCATCGACTCTGCTGAGAAATACACACAGAGTGGGAGACAAA 3843
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3844 GGGACTCAGCTGTCCGGCGGCGAGAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGA 3903
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3904 CAGGCTTCATCTTACTTCTGGATGAAGCGCATCAGCTCTGGATACGAGAGTGAAG 3963
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3964 GTCGTCCAGGAAGCGCTGCAACAGCCAGGAGGCCCGCCTGCAATGTGATCGGCAC 4023
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4024 CGCCTGTCCAGCATCCAGAGCCAGACTTGATCGTGTGATTCAGAACCGCGCAGGTCAAG 4083
```

```
QY 788 GluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4084 GAGCAGCGCACCCAGCAGCTGCTGGCCCAAGAGGCATCTATTTCTCGATGGTCACT 4143
QY 808 AlaGlnSer 810
Db 4144 GTGCAGGCT 4152
RESULT 9
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MOR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1
Alignment Scores:
Pred. No.: 3,47e-230 Length: 4317
Score: 2438.00 Matches: 467
Percent Similarity: 76.24% Conservative: 162
Best Local Similarity: 56.61% Mismatches: 178
Query Match: 59.77% Indels: 18
DB: 9 Gaps: 4
US-09-873-409-2 (1-812) x US-10-044-671-1 (1-4317)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1429 ATTGATGGACAGGACATTTAGGACCAATAATGTAAGGCATCTTCGGGAATATTACTGGTGTG 1488
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1489 GTGAGTCAGAGGCTGTGTGTTTGGCCACCACCATAGCTCAAAACATTCGCTATGGCCGC 1548
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1549 GAAATATCTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCAATGCTATGATTTT 1608
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1609 ATCATGAAACTACTATAATAATTTGACACTCTGGTTGGAGAGAGAGGGGGCCGCTGAGT 1668
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1669 GGTGGACAGAAACAGAGATCGCCATTTGCTCGGGCCCTGGTTCCGAACCCAGATTTCTT 1728
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1729 CTGCTGATGAGGCAACGTCAGCTCTGACACTCTGAAACACTGAAAGTGAAGCAGTGGTTCAAGTGGCC 1788
QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1789 CTGGATAAGGCCAGAAAGGCGGAGCTTACCATTTGTGTAGTACTATCATCTTTGTCTACAGTT 1848
```

Qy	142	ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis	161
Db	1849	CGTAATCCGATGTCATTTGCTGTTTGGATGATGAGTCATTTGGAGAAAGGAATCAT	1908
Qy	162	AlaLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys	181
Db	1909	GATGAATCATGAAGAGAGGCGATTTACTTCAACATTTGTCACATCGC---ACAAG	1965
Qy	182	LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu	201
Db	1966	GGAAATCGAAATTCAGTTAGAAATGCCATCGCTGAATCCAAAAGTGAAGTGATGCCTTG	2025
Qy	202	ProLeuHisSerValLysSerIleLysSerAspPheIle-----	214
Db	2026	GAATG---TCTCAAAGATTTCAGGCTCCAGTTTAATAAAAAGAGATCAACTCGCAGG	2082
Qy	215	-----AspLysAlaGluSerThrGlnSerLysGluIle	226
Db	2083	AGTATACATCCACACAGCCCAAGACAGAAAGCTTGGTACAAAAGAGAGCTTGAATGAG	2142
Qy	227	SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPhe	246
Db	2143	AATGTTCTTCAGTTTCTCTCGAGGATTCGAAGCTGAATCAACTCAATGATGCCTTAT	2202
Qy	247	ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle	266
Db	2203	TTTGTTGGTTGGTATATTTTGTCATTTAAMACGAGCGCTGCACACAGCATTTTCAATA	2262
Qy	267	IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis	285
Db	2263	ATATTTCAAGGATATAGGATCTTTACCAGATGAGGATCCTGAAACAAACAGCAG	2322
Qy	286	AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr	305
Db	2323	AATGATAACATGTTTCTGTATTGTTCTAGTCTTGGAAATTAATTTCTTTTATTACATTT	2382
Qy	306	PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis	325
Db	2383	TTCTCTCAGGGCTTCACATTTTGCAAGCTGGGAGATCTCTACTAAGCGCTTCGATAC	2442
Qy	326	LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSer	345
Db	2443	ATGGTTTTTCAGATCCATGCTGACACAGATGTCAGCTGGTTGATGACCCCTAAAAACACC	2502
Qy	346	ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly	365
Db	2503	ACTGGACATTGCAACACCGGCTTGCCAATGATCGGGCTCAAGTTAAAGGGGCTATAGGT	2562
Qy	366	SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer	385
Db	2563	TCCAGGCTTGCTGTCATTACCCAGAATATAGCAATCTTGGACAGCGCATTTATTATATPCC	2622
Qy	386	PheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal	405
Db	2623	TTAATCTATGTTGGCAATTAAACACTTTTACTCTTAGCAATTTGTACCCATCATTTGCAATA	2682
Qy	406	ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu	425
Db	2683	GCAGGAGTTGTTGAAATGAAAATGTTGTTCTGCGACAGCACTGAAAGATAAAGAAAGAGCTA	2742
Qy	426	LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu	445
Db	2743	GAAGAGCTGGGAAGATTGCTACGAAGCGATCGAAAACCTCCGAACCTGTGTTTCTTTG	2802
Qy	446	ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetGlnThrGlnHisArgAsn	465
Db	2803	ACTCGGAGCAGAAGTTTGAATACATGATATGCACAGAGTTTGCAGAGTACCATACAGAAAC	2862
Qy	466	ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr	485
Db	2863	TCCTTTAGGAAAGACACATCTTCGGGGTCTCATCTTTCTATCACCCAGCAATGATGTAT	2922

Qy	486	Phe	Ala	Tyr	Ala	Ala	Gly	Phe	Arg	Phe	Gly	Ala	Tyr	Leu	Ile	Gln	Ala	Gly	Arg	Met	Thr	505
Db	2923	TTTTCTTAT	CTCTGCTGTTT	TCGGT	TTGCTGCTACT	TTTGGT	GGCAAA	TGAGT	TTCA	GAAC	2982											
Qy	506	Pro	Glu	Gly	Met	Phe	Ile	Val	Phe	Thr	Ala	Ile	Ala	Tyr	Gly	Ala	Met	Ala	Ile	Gly	525	
Db	2983	TTTC	CGA	GTG	TTCTTT	TGGTAT	TCTCAG	CTAT	TGCTTT	GTG	TCCAT	TGGC	ATG	GGC	AGT	GGG	CAG	3042				
Qy	526	Thr	Leu	Val	Leu	Ala	Pro	Glu	Tyr	Ser	Leu	Ser	Ala	Lys	Ser	Gly	Ala	Ala	His	Leu	545	
Db	3043	GTC	AGT	TTCA	TTGCT	CTG	ACTAT	GC	CAAG	CCAA	GAAG	TAT	CAG	CA	GC	CCAC	GC	3102				
Qy	546	Leu	Leu	Glu	Lys	Pro	Asn	Ile	Asp	Ser	Arg	Ser	Gln	Glu	Gly	Ser	Pro	Asp	Thr	565		
Db	3103	ATCA	TTG	AAAA	AAAG	CCCT	CTGA	TTG	AC	TG	AC	TG	AC	CC	CTC	AG	CC	3162				
Qy	566	Cys	Glu	Gly	Asn	Leu	Glu	Phe	Arg	Glu	Val	Ser	Phe	Thr	Pro	Cys	Arg	Pro	Asp	Val	585	
Db	3163	TTGG	AGGA	AA	TTG	ACAT	TTA	TGAG	TCG	TGTT	CA	ACT	TAT	CC	CA	CT	CG	3222				
Qy	586	Phe	Ile	Leu	Arg	Gly	Leu	Ser	Leu	Ser	Ile	Glu	Arg	Gly	Lys	Thr	Val	Ala	Phe	Val	605	
Db	3223	CCC	TG	CTC	AGG	GGT	TG	AG	CC	CTC	CG	AG	TG	AA	AG	GG	CG	3282				
Qy	606	Ser	Ser	Gly	Cys	Gly	Lys	Ser	Thr	Ser	Val	Gln	Leu	Gln	Arg	Leu	Tyr	Asp	Pro	Val	625	
Db	3283	AGC	AGT	GG	CTG	TGG	AG	AG	CA	CA	G	AGT	GT	TG	CT	CA	CT	3342				
Qy	626	Gln	Gly	Gln	Val	Leu	Phe	Asp	Gly	Val	Asp	Ala	Lys	Glu	Leu	Asn	Val	Ile	Trp	Leu	645	
Db	3343	GCT	GGT	TCA	GTG	CTA	TTG	ATG	TGC	AA	AG	AT	AA	AG	CA	CT	TG	3402				
Qy	646	Ser	Gln	Ile	Ala	Ile	Val	Pro	Gln	Glu	Pro	Val	Leu	Phe	Asn	Cys	Ser	Ile	Ala	Glu	665	
Db	3403	GC	AC	CA	CT	GG	CG	CA	CT	CG	T	CT	CA	AG	AG	CC	CT	3462				
Qy	666	Ile	Ala	Tyr	Gly	Asp	Asn	Ser	Arg	Val	Val	Pro	Leu	Asp	Glu	Ile	Lys	Glu	Ala	Ala	685	
Db	3463	ATT	GC	CTAT	GG	AG	CA	AC	CC	GGG	TG	TAT	CA	CA	TG	AA	AG	3522				
Qy	686	Ala	Ala	Asn	Ile	His	Ser	Phe	Ile	Glu	Gly	Leu	Pro	Glu	Lys	Tyr	Asn	Thr	Gln	Val	705	
Db	3523	GAG	CCA	CA	AT	CA	CA	CT	TT	CA	TC	AG	CA	CA	CT	CC	CT	3582				
Qy	706	Leu	Lys	Gly	Ala	Gln	Leu	Ser	Gly	Gln	Lys	Gln	Arg	Leu	Ala	Ile	Ala	Arg	Ala	Leu	725	
Db	3583	GACA	AA	GG	AA	CC	CA	GT	CT	CT	GT	GG	TG	CC	AG	AA	CA	3642				
Qy	726	Leu	Gln	Lys	Pro	Lys	Ile	Leu	Leu	Leu	Asp	Glu	Ala	Thr	Ser	Ala	Leu	Asp	Asn	Asp	745	
Db	3643	GTT	AG	CA	GC	CT	CA	TAT	TTT	TG	CTTT	TG	GA	TG	AA	AG	AT	3702				
Qy	746	Glu	Lys	Val	Val	Gln	His	Ala	Leu	Asp	Lys	Ala	Arg	Thr	Gly	Arg	Thr	Cys	Leu	Val	765	
Db	3703	GA	AA	GG	TTG	TCCA	AG	AG	CC	CT	TGG	CA	AA	AG	CC	AG	AA	3762				
Qy	766	Thr	His	Arg	Leu	Ser	Ala	Ile	Gln													

```

; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR FILING DATE: 09/584,586
; PRIOR APPLICATION NUMBER: 2000-05-31
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-866-866A-1

Alignment Scores:
Pred. No.: 4,52e-230 Length: 3860
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x US-09-866-866A-1 (1-3860)

Qy 2 ValAspGluAenAspIleAArgAlaLeuAenValAArgHisIleGlyVal 21
Db 1357 GTTCATGGACAGATATTAGACCATTAATGTAAGTTCTACGGGAATCATTTGGTGTG 1416

Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAenAenIleLysTyrGlyArg 41
Db 1417 GTGAGTCAGGAACCTGTATTGTTGCCACCACGATAGTCGTAAGAAACATTCGCTATGGCGGT 1476

Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe 61
Db 1477 GAAATGTCACATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCCATGACTTT 1536

Qy 62 IleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1537 ATCATGAAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGAGGGCCCATTTGAGT 1596

Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValAArgAenProLysIleLeu 101
Db 1597 GGTGGCAGAGACGAGATCGCCATTCGACGTGCCCTGTTGCAACCCCAAGATCCTC 1656

Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1657 CTGCTGGATGAGGCACGTCAGCCTTGACACAGAAAGCAGAGAGTGTTCAGGTGGCT 1716

Qy 122 LeuGlyLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1717 CTGATAGGCCAGAAAGGTGCGACACCATTTGTAGTGTCTATGTTGCTACAGTT 1776

Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1777 CGTAATGTCAGCTCATCGCTGTTTCGATGATGAGTCAATTTGGGCAAGAAAGAAATCAT 1836

Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeuValMetSerGln----- 178
Db 1837 GATGAACATCATGAAGAAAGGAGGATTTACTTCAAACTTGTCAAACTGCACAGACGAGGA 1896

Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 1897 AATGAAGTTGAATTAGAAAATGACGCTGATGAATCCAAAGTGAATTTGATGCTTGGAA 1956

Qy 196 ArgLysThrAenSerLeuProLeuHisSerVal-----LysSerIle 209
Db 1957 ATGCTTCAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAAACTCGTAGGAGTGTG 2016

```

```

Qy 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 2017 CGTGGATCAACAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGTATA 2076

Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGluTrpProPheValVal 248
Db 2077 CTTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCCTTTATTTGT 2136

Qy 249 LeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPheSerIlePhe 268
Db 2137 GTTGGTGTATTGTCGCAATTATAAATGGAGCCTGCAACAGCATTGTCAAATAATATT 2196

Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2197 TCAAGATTATAGGGTTTTTACAAGATTGATGATCCTCAACAAACAAACACAGATAGT 2256

Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 2257 AACTTGTGTTTCACTATTGTTCTAGCCCTTGAATATTATTTCTTTTATCATTTTCTCTT 2316

Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2317 CAAGGTTTTCATTTGGCAAAGCTGGAGATCCTCACCAGCGGCTCCGATACATGTT 2376

Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAenSerThrGly 347
Db 2377 TTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAACACACCATGGA 2436

Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2437 GCATTGACTACCAAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAG 2496

Qy 368 IleGlyValLeuThrGlnAenAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2497 CTTGCTGTAATACCCAGAAATATAGCAAACTTTGGGACAGAAATAATATCTCTTCATC 2556

Qy 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2557 TATGTTGGCACTACACTGTTACTTTAGCAATTTGTACCCATCATTTGCAATAGCAGGA 2616

Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAenLysAspLysGlnGluLeuLysHis 427
Db 2617 GTTGTGAAATGAAATGTTGCTCGACAAAGCTGAAAGATGAAGAAAGCACTAGAGGT 2676

Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIleValSerLeuThrArg 447
Db 2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGAACCGTTGTTCTTTGACTCAG 2736

Qy 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2737 GAGCAGAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAACTCTTTG 2796

Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2797 AGGAAAGCACACATCTTTGGAAATTACATTTCTTCAACCCAGGCAATGATGATATTTTCC 2856

Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2857 TATGCTGGATGTTTCGGGTTTGGAGCCTACTTGGTGCAATAACTCATGAGCTTTGAG 2916

Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 2917 GATGTTCTGTTAGTATTTCAGCTGTGTTCTTTGGTGCCATGGCCGTGGGCAAGTCAGT 2976

Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 2977 TCATTTGCTCTGACTATGCAAAAGCCAAATATATCAGCAGCCCAATCATCATGATCAT 3036

Qy 548 GluLysLysProAenIleAspSerArgSerGlnGluGlyLysProAspThrCysGlu 567
Db 3037 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGCCCTAATGCGCAACACATTTGAA 3096

```

```
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrPrProCysArgProAspValPheIle 587
Db 3097 GGAATGTCACATTTGGTGAAGTTGTTAACTTATCCACCCGACCGGACATCCCGAGT 3156

QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3157 CTTGAGGAGTGAAGCTGGAGAGTGAAGAGGGCCAGACGCTGGCTCTGGTGGGAGGAGT 3216

QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3217 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCTGGAGCGGTCTTACGACCCCTTTGGCAGG 3276

QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGln 647
Db 3277 AAAGTGTCTGTTGATGGCAAAAGAAATAAGACGATGAATGTCAGTGGCTCCGAGCACAC 3336

QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3337 CTGGGATCGTGTCCAGAGAGCCCATCTCTGTTGACTGCGAGCATGCTGAGAACATTGCC 3396

QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3397 TATGGAGACAACAGCCGGTGGTGTCTACAGGAAGAGATCGTGAGGGCAGCAAGAGGCC 3456

QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3457 AACATACATCGCTTCATCGAGTCACTCGCTTAATAAATATAGCACATAAGTAGGAGACAAA 3516

QY 708 GlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 3517 GGAACTCAGCTCTCTGTGTGGCCGAAACACAGCATGCCATGCTGCGCTTGTGTAGA 3576

QY 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3577 CAGCTCATATTTGCTTTTGGATGAGCACCGTCAGCTCTGGATCAGAAAGTGAAG 3636

QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3637 GTTGTCCAGAAGCCCTGGCAAAAGCCAGACAGTGAATAGTGTGTTCAGAAATGCGAGTCAAG 3696

QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 3697 CGCTGTGCCACCATCCAGATGACAGCTTAATAGTGTGTTCAGAAATGCGAGTCAAG 3756

QY 788 GlyGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3757 GAGCATGGCAGCATGACAGTCTGCTGGCAGACAGAAAGGCATCTATTTTCAATGGTCACT 3816

QY 808 AlaGlnSer 810
Db 3817 GTCAGGCT 3825

RESULT 11
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
```

```
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Pred. No.: 4,62e-230 Length: 3860
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x US-09-866-866A-3 (1-3860)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1357 GTTGATGGACAGATATTAGGACCATAAATGTAAGGTTCTTACGGGAAATCATTTGGTGTG 1416

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1417 GTGAGTCAGGAACCTGTATTGTTTGGCCACCATAGCTGAAACACATTGCTATGGCCGT 1476

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1477 GAAATGTGTCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCTATGACTTT 1536

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1537 ATCATGAAACTGCTCATAAATTTGACACCTGTTGGAGAGAGGGGCCAGTTGAGT 1596

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1597 GGTGGCAGAAAGCAGAGGATCGCCATTGTCAGCTGCTGTTCCCAACCCCAAGATCCTC 1656

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCGTGTTCAGTGGCT 1716

QY 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 1717 CTGGATAAGCCAGAAAGGTCGACCAACCATTTGTATAGTCTCATCGTTTGTCTACAGTT 1776

QY 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 1777 CGTAATGCTGACGTCATCGCTGCTGTTTCGATGATGAGTCAATTGTGGAGAAAGAAATCAT 1836

QY 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 1837 GATGAACCTCATGAAGAGAAAGGACATTACTTCAAACCTTGTTCACAATGCACAGCAGGA 1896

QY 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 1897 AATGAAGTTGAATTAGAAATGCGCTGATGATGATCCAAAGTGAATGATGCTTGGAA 1956

QY 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 1957 ATGCTCTCAAATGATTCAAGATCCAGTCTTAATAAGAAAAAGATCAACTCGTAGGAGTGC 2016

QY 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228
Db 2017 CGTGGATCAACGCCCAAGACAGAAAGCTAGTACCAAGAGGCTCTGGATGAAGATATA 2076

QY 229 ProGluValSerLeuLysIleLeuLysAsnLysProGluTrpProPheValVal 248
Db 2077 CTTCCAGTTTCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTGTTT 2136

QY 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 2137 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCTGCAACACAGCAATTGCAATAATATT 2196

QY 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 2197 TCAAGATTATAGGGGTTTTTTACAGAATTGATGATCTCTGAAACAAACACGACGAATAGT 2256
```

```
QY 288 GluileTyrSerMetIlePheValIleuGlyValIleCysPheValSerTyrPheMet 307
Db 2257 AACTGTGTTTCTACATATGTTTCTAGCCCTTGGAAATATATTTCTTTATACATTTTTCCTT 2316
QY 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 2317 CAAGGTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGCTCCGATACATCGTTT 2376
QY 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlnAsnSerThrGly 347
Db 2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCTCAAAAACACCCACTGGA 2436
QY 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 2437 GCATTTGACTACAGGCTCGCCCAATGATGCTCAGTTAAAGGGGCTATAGTTTCAGG 2496
QY 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db 2497 CTTGCTGTAAATTACCCAGAAATATAGCAATCTTGGGACAGGAATAATATATCTTCTCATC 2556
QY 388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 2557 TATGGTTGGCAACTAACACTGTTACTCTAGCAATTTGACCCATCTTCGCAATAGCAGGA 2616
QY 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 2617 GTTGTGTAATGAATAATGTTGCTCGGACAGCACTGAAGATAAGAAAGAACTAGAAAGT 2676
QY 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAACTTCCGAACCGTTGTTCTTTGACTCAG 2736
QY 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 2737 GAGCAGAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTTG 2796
QY 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 2797 AGGAAAGCACACATCTTTGGAAATATACATTTTCTTCCACCCAGGCAATGATATTTTTC 2856
QY 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 2857 TATGCTGATCTTTCCGGTTTGGAGCTTACTTGTGGCACATAAATCATGAGCTTTTGAG 2916
QY 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlySerThrLeu 527
Db 2917 GATGTTCTGTAGTATTTTTCAGCTGTTGCTCTTGTGTCATGGCCGTCGGGCAAGTCAGT 2976
QY 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 2977 TCATTTGCTCTGCTATGCCAAGCCAAATATATACAGAGCCCATCATCATCATCATCAT 3036
QY 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 3037 GAAAAAACCCCTTTGATTGACAGCTACAGCAGGAGGCCCTTAATGTCGCAACACATTTGAA 3096
QY 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 3097 GGAATATGTACATTTTGTGTGAAGTTGTATTCAACTATCCACCCGACCGGACATCCAGTG 3156
QY 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 3157 CTTGAGGACTGAGCCTGGAGGTGAAGAGGGCCAGACGCTGGCTCTGCTGGGCGAGCAGT 3216
QY 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 3217 GCCTGTGGAAGAGACACAGTGTCCAGTCTCTGGAGCGGTTCTACGACCCCTTGGCAGG 3276
QY 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db 3277 AAAGTGTCTGTGTCGCAAGAAATAAAGCACTGAATGTTTCAGTTGCTGCTCCGAGCACAC 3336
```

```
QY 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 3337 CTGGGCATCGTGTCCAGGAGCCCATCTGTTTGGACTGTGAGCATTTGCTGAGAACATTGCC 3396
QY 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 3397 TATGGAGACAACAGCCGGTGTGTCTACAGGAAGAGATCGTGGAGGCAGCAAGGAGGCC 3456
QY 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 3457 AACATACATGCTTTCATCAGTCACTGCTCTTAATATATAGCACTAAAGTAGGAGACAAA 3516
QY 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727
Db 3517 GGAACCTCAGCTCTCTGTTGGCCAGAAACAACGCAATTGCCATAGCTCGTGCCCTTGTAGA 3576
QY 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 3577 CAGCCTCATATTTTGTCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAG 3636
QY 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 3637 GTTGTCAAGAACCTTGGACAAGCCAGAGAAGCCGACCTGCACTTGATGTGCTCAC 3696
QY 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLys 787
Db 3697 CGCCTGTCCACCATTCCAGAAATGCACTTAATAGTGTGTTTCAGAAATGCGAGTCAAG 3756
QY 788 GluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 3757 GAGCATGGCAGCATCAGCAGCTGTGGCACAGAAAGGCAATCTATTTTCAATGGTCAGT 3816
QY 808 AlaGlnSer 810
Db 3817 GTCCAGGCT 3825
RESULT 12
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US0020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2
Alignment Scores:
Pred. No.: 6,08e-230 Length: 4643
Score: 2436.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 9 Gaps: 4
US-09-873-409-2 (1-812) x US-10-072-621-2 (1-4643)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1778 GTTGATGGACAGGATATTAGACCATTAATGTAAAGTTTCTACGGGAATCATATTGGTGTG 1837
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
```



Db 1838 GTAGTCAGGAACCTGTATTGTTGGCCACACGATAGCTGAAAACATTCCTGCTATGGCCGT 1897  
Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
Db 1898 GAAAAATGTCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCATGCTATGACTTT 1957  
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
Db 1958 ATCATGAACTGCTCATAAATTTGACACCTGTGTCAGAGAGAGAGGGGCCAGTTGAGT 2017  
Qy 82 GlyGlyGlnLysGlnArgGlyAlaAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101  
Db 2018 GGTGGGCAAGACAGAGATCGCCATTCACAGTCGCTGTTGTCACACCCCAAGATCCCT 2077  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
Db 2078 CTGCTGATGAGCCACGTCAGCCTTCGACACAGAAAGCAGAGTGGTTCCAGTGGCT 2137  
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 2138 CTGATAAGGCCAGAAAGGTGCGACCACTTGTGTAGTCTCATCTGTTGTCTACAGTT 2197  
Qy 142 ArgSerAlaAlaLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
Db 2198 CGTAATCTGACGTCATCGCTGTTTCGATGATGAGTCAATTTGGAGAAAGGAATCAT 2257  
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
Db 2258 GATGAACTCATGAAGAGAAAGGCAATTTACTTCAAACTTGTCAATGTCAGACAGCAGGA 2317  
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
Db 2318 AATGAAGTTGAAATTAGAAAATGACGCTGATGAAATCCAAAGTGAAATGATGCTTGAA 2377  
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
Db 2378 ATGTCTTCAAAATGATTCAGATCCAGTCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGC 2437  
Qy 210 Lys---SerAspPheIleAspLysAlaGluLysSerThrGlnSerLysGluIleSerLeu 228  
Db 2438 CTGTGGATCAACGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAAGTATA 2497  
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluIrrProPheValVal 248  
Db 2498 CTTCCAGTTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCTTATTTGTT 2557  
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
Db 2558 GTTGGTGATTTTGTGCCATTTATAAATGGAGGCTGCAACCGCATTTGCAATAATATT 2617  
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287  
Db 2618 TCAAGATTATAGGGTTTTTACAGAAATTTGATGATCTCGAACAACCAACAGCAATAGT 2677  
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307  
Db 2678 AACTTGTTTCTACTATTGTTCTAGCCCTTGAATTTCTTTTATTATATTTTCTT 2737  
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAla 327  
Db 2738 CAGGGTTTCATTTGGCAAGCTGGAGAGATCTCTCAACGAGGCTCCGATATCATGGT 2797  
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrrPheAspGluLysGluAsnSerThrGly 347  
Db 2798 TTCCGATCCATGTCACAGAGATGAGTGGTTGATGATCCCTAAACACCACTGGA 2857  
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367  
Db 2858 GCATTGACTACAGGCTCGCAATGATGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2917  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 2918 CTTGCTGTAATTACCCAGAAATATAGCAAAATCTTGGGACAGGAATAATATATATCTTCATC 2977

Qy 388 TyrGlyTrrPgluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 2978 TATGGTTGGCAACTACACTGTACTTCTTAGCAATTTGTACCCATCATTTGCAATAGCAGGA 3037  
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
Db 3038 GTTGTGTAATGAAATGTTGTCGCAAGCAGCACTGAAAGATAAGAAAGAACTAGAGGT 3097  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
Db 3098 GCTGGGAAGATCGCTACTGAAGCAATAGAAAATTCGGAACCGTTGTTCTTTGACTCAG 3157  
Qy 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
Db 3158 GAGCAGAAGTTTGAACATATGATGCTCAGAGTTTTCAGGTACCATACAGAAACTCTTTG 3217  
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 3218 AGAAAGCACACATCTTTGGAATTTACATTTCTTCCACCGAGCAATGATGTATTTTCC 3277  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 3278 TATGCTGGATGTTTCGGTTTGGAGCTACTTGGTGGCACATAAACTCATGACTTTGAG 3337  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
Db 3338 GATGTTCTGTGTAGTATTTTTCAGCTGTTGTTGTTGGTGCCATGCGCGGCAAGTCAGT 3397  
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
Db 3398 TCATTTTCTCTGACTATGCCAAAGCCAAATATATACGACAGCCACATCATGATCAT 3457  
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567  
Db 3458 GAAAAAACCCCTTTGATTTGACAGCTACAGCAGGAAGGCTTAATGCCGAACATTTGGAA 3517  
Qy 568 GlyAsnLeuGluPheArgGluValSerPheTyrProCysArgProAspValPheIle 587  
Db 3518 GGAATGTCACATTTTGTGTAAGTTGTATTCAACTATCCACCCGACCGACATCCCATG 3577  
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgLysThrValAlaPheValGlySerSer 607  
Db 3578 CTTTCAGGATGAGCTGAGCTGGAGGTGAAGAGGGCCACGCTGCTCTGTGGGACGAGT 3637  
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627  
Db 3638 GGTGTGGGAAGACACAGTGTCTCAGCTCTGAGCGGTTCTACGACCCCTTGGCAGGG 3697  
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnIrrPheArgSerGln 647  
Db 3698 AAGTGTCTGCTGATGCAAGAAAGAAATAAAGCGACTGAATGTTTCAGTGGCTCCGACAC 3757  
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667  
Db 3758 CTGGGATCGTGTCCCGAGGAGCCCATCTCTTTGACTGCGACATTTGCTGAGAACATTTGC 3817  
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluLysGluAlaAlaAsnAlaAla 687  
Db 3818 TATGGAGAACAAAGCCGGGTGGTGTACAGGAAGAGATTGTGAGGGGAGCAAGAGAGGCC 3877  
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707  
Db 3878 AACATACATCGCTTCATCGAGTCACCTGCCTTAATAATATAGCACTAAAGTAGGAGACAA 3937  
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727  
Db 3938 GGAACCTCAGCTCTCTGCTGGCCAGAAACAAACGATTTGCCATGTCTGCTGCTTGTAGA 3997  
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaIleuAspAsnAspSerGluLys 747  
Db 3998 CAGCCTCATATTTTGTCTTTTGGATGAAGCCACGCTGAGCTCTGATACAGAAAGTAAAG 4057



```
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4058 GTTCTCCAGAGAGCCCTGGCAGAGCCAGAGAGCGCCACCTGCTGATTGTGCTCAC 4117
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4118 CGCCTGTCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGAAATGCAGAGTCAAG 4177
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4178 GAGTCGCGCATCAGCAGCTGCTGGCAGAGAGGACTTAATTTTCAATGCTCAGT 4237
Qy 808 AlaGlnSer 810
Db 4238 GTCCAGGCT 4246

RESULT 13
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306.417
; EARLIER FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SPbeta71m4
; US-09-306-417-1

Alignment Scores:
Pred. No.: 1.52e-229 Length: 8630
Score: 2435.00 Matches: 466
Percent Similarity: 76.79% Conservative: 166
Best Local Similarity: 56.62% Mismatches: 177
Query Match: 59.72% Indels: 14
DB: 10 Gaps: 4
```

```
US-09-873-409-2 (1-812) x US-09-306-417-1 (1-8630)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 2576 GTTGTGGACAGATATTAGGACCATAAATGTAGATTCTACGGGAATCATTTGGTGTG 2635
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 2636 GTGAGTCAGAGAACCTGTAATTGTTTGCACCACGATAGTGAAACACATTCGCTATGCGCT 2695
Qy 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 2696 GAAATATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAGCCCAATGCTATGACTTT 2755
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 2756 ATCATGAACCTGCTCATAAATTTGACCCCTGTTGGAGAGAGGGGCCCATGTGAGT 2815
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2816 GTTGGCAGAGCAGAGGATCGCCATTGCACGTGCTGTTTCGCAACCCCAAGATCCTC 2875
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 2876 CTGCTGGATGAGGCCACGTCAGCCTTTGGACACAGAAAGCGAGCAGTGGTTCAAGTGGCT 2935
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db 2936 CTGGAAGGCCAGAAAGTCGAGCCACCATTTGTATAGTCACTGTTTGTCTACAGTT 2995
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db 2996 CGTAATGCTGACCTCATCGCTGCTTTCGATGATGGAGTCATTGTGGAGAAAGAAATCAT 3055
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178
Db 3056 GATGAACCTCATGAAGAGAAAGGCACTTTACTTCAAACTTGTTCACAAATGCGACAGCAGCA 3115
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db 3116 AATGAAGTTGAATAGAAAATGCGCTGATGAATCCAAAGTGCATAATGATGATGCTTGGAA 3175
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209
Db 3176 ATGTCCTTCAAAATGATTCACAGATCCAGTCTAATAAGAAAAGATCAACTCGTAGGAGTGC 3235
Qy 210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db 3236 CGTGGATCAACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAAGTATA 3295
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db 3296 CCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTAATACTGAATGGCTTATTTGTT 3355
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db 3356 GTTGTGTATTGTCGCATTATAAATGGAGGCTGCAACACGAGCATTTGCAATAATATT 3415
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db 3416 TCAAGATTATAGGGGTTTTTACAGAAATTTGATGATCCTCGAAACAAACACAGAAATAGT 3475
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db 3476 AACTTGTGTTTCACTATTGTTCTAGCCCTTGGAAATTAATTTCTTTTATACATTTTCTT 3535
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db 3536 CAAGTTTTCATTTTGGCAAGCTGGAGAGATCTCTACCAAGCGGCTCCGATACATGTT 3595
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db 3476 AACTTGTGTTTCACTATTGTTCTAGCCCTTGGAAATTAATTTCTTTTATACATTTTCTT
```

```
Db 3596 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAACACCACTCGGA 3655
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db 3656 GCATTGACTACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 3715
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleSerPheIle 387
Db 3716 CTTGCTGTAATACCCAGAAATATAGCAAACTCTGGGACAGCAATAATATATCTTCATC 3775
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db 3776 TAGGGTTGGCACTAACACCTGTACTCTTAGCAATGTACCCATCATTCGCAATAGCAGGA 3835
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db 3836 GTTGTTGAATGAATGTTGCTCGACACAGCACTGAAGCAATAAGAAAGAACTAGAGGT 3895
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db 3896 GCTGGGAAGATCGTACTGAAGCAATAGAAAACCTTCGGAACCGTGTCTTCTTCACTCAG 3955
Qy 448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db 3956 GAGCAGAAGTTTGAACATATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTG 4015
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db 4016 AGAAAGCACACATCTTTGGAAATACATTTCTTCCACCGAGCAATGATGATTTTCC 4075
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db 4076 TATGCTGGATGTTTCCGGTTCGAGCTACTTGGTGACATATAAATCAATCATGAGCTTGAG 4135
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db 4136 GATGTTCTGTAGTATTTTTCAGCTGTTGTTCTTGGTGCCATGCGCGTGGGCAAGTCAGT 4195
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db 4196 TCATTGCTCTGACTATGCAAGCCCAAAATATATCAGCAGCCCATCATCATGATCAT 4255
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db 4256 GAAAAACCCCTTTGATTGACAGCTACAGCAGGAGCGCTTAATGCGCAACACATTTGAA 4315
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db 4316 GGAATGTCACATTTGTTGAAGTTGTATTCAACTATCCACCCGACCGGACATCCCAATG 4375
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db 4376 CTTCAGGGACTGAGCCTGGAGGTGAAGAGCGCCAGACGCTGGCTCTGGTGGGCGAGCAGT 4435
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db 4436 GGCTGTGGGAAGACACACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGACCCCTTGGCAGG 4495
Qy 628 GlnValLeuPheAspGlyValAlaAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGln 647
Db 4496 AAAGTGTGCTGTGATGGCAAAAGAAATTAAGCGACTGAATGTTCACTGGCTCCGAGCAC 4555
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db 4556 CTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGACGATGCTGAGAACTTGCC 4615
Qy 668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db 4616 TATGAGACACACACCGGGTGGTGTACAGGAAGAGATCGTGGAGGCGCAAGAGAGGCC 4675
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db 4676 AACATACATGCTTTCATCGAGTCACTGCTTAATAATATAGCACTAAAGTAGGAGACAAA 4735
```

```
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db 4736 GGAACCTCAGCTCTCTGTTGGCGCAGAAACAACGCAATTCGCCATAGCTCGTCCCTTGTAGA 4795
Qy 728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db 4796 CAGCCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACAGAAAGTGAAGAG 4855
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db 4856 GTTGTCCAGAAGCCCTGCAAAAGCCAGCAAGAGCGCCGCACTGTGATTTGATGCTCAC 4915
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db 4916 CGCCTCTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTCAGAAATGCGCAGTCAAG 4975
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db 4976 GAGCATGGCCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 5035
Qy 808 AlaGlnSer 810
Db 5036 GTCCAGGCT 5044

RESULT 14
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20030103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta91msA1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: msA1 mdrl1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
```

US-09-306-417-2

## Alignment Scores:

Pred. No.: 1.52e-229 Length: 8630  
Score: 2436.00 Matches: 466  
Percent Similarity: 76.79% Conservative: 166  
Best Local Similarity: 56.62% Mismatches: 177  
Query Match: 59.72% Indels: 14  
DB: 10 Gaps: 4

US-09-873-409-2 (1-812) x US-09-306-417-2 (1-8630)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
Db 2576 GTTGATGACAGGATATTAGGACCATAAATGTAAGGTTCTACGGGAATCATTTGGTGTG 2635  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
Db 2636 GTGAGTCAAGAACCTGTATTGTTGCCACCACGATAGCTGAANAACATTCGCTATGGCCGT 2695  
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
Db 2696 GAAATGTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2755  
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
Db 2756 ATCATGAACCTGCCTCATAAATTTGACACCTGTTGAGAGAGAGGGCCGCTGAGT 2815  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
Db 2816 GGTGGGCAAGACAGAGATCGCCATTCGATGCGTCCCTGTTGCAACCCCAAGATCCTC 2875  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
Db 2876 CTGCTGATGAGGCACGCTGACCTTGACACAGAAAGCGAGCAGTGTTCAGTGGCT 2935  
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 2936 CTGGATAAGGCAGAAAAGGTCGACACCACTTGTGATGATGATGATGATGATGATGAT 2995  
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
Db 2996 CGTAATGCTGACGTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3055  
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeuValMetSerGln----- 178  
Db 3056 GATGAACCTCATGAAGAAAGAGGATTTACTTCAAACTTGTCACAAATCCACACGACG 3115  
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195  
Db 3116 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATGATGCTTGAA 3175  
Qy 196 ArgLysThrAsnSerLeuProLeuHisSerVal-----LysSerIle 209  
Db 3176 ATGCTTCAAAATGATTCAGATCCAGTCTAATAAGAAAAGATCAACTCGTAGGAGTGC 3235  
Qy 210 Lys---SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228  
Db 3236 CGTGATCAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGTATA 3295  
Qy 229 ProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248  
Db 3296 CTTCCAGTTTCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGGCTTATTTTGT 3355  
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIlePhe 268  
Db 3356 GTTGGTGTATTGTGTCATTTAATAATGGAGCCCTGCAACACGACGATTTGCAATAATTT 3415  
Qy 269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAla 287  
Db 3416 TCAAGATATTAGGGTTTTTACAGAAATGATGATCTCTGAAACAAACGACAGAAATAGT 3475  
Qy 288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307

Db 3476 AACTGTGTTTCACTATTGTTTCTAGCCCTTGGAAATATTCTTTTATCATTTTCCCT 3535  
Qy 308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327  
Db 3536 CAAGGTTTCACATTTGGCAAGCTGGAGATCTCTACCAAGGGCTCCCATCATGTT 3595  
Qy 328 PheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGly 347  
Db 3596 TTCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCTAAAAACACCACCTGGA 3655  
Qy 348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnGlnGlyAlaThrGlySerArg 367  
Db 3656 GCATTCGACTACCAAGCTCGCAATGATGCTGCTCAAGTTAAAGGGGTATAGGTTCAGG 3715  
Qy 368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387  
Db 3716 CTTGCTGTAATACCCAGAAATATAGCAATCTTTGGGACAGAAATATATATCTTCTATC 3775  
Qy 388 TyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407  
Db 3776 TATGTTTGGCAACTAACACATGTTTACTCTTAGCAATGTACCCATCATTTGCAATAGCAGGA 3835  
Qy 408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427  
Db 3836 GTTGTGAAATGAAATGTTTGTCTGCAAGCACTGAAAGATAGAAAGAACTAGAAAGGT 3895  
Qy 428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447  
Db 3896 GCTGGGAAGATCGTACTGAAGCAATAGAAAATCTCGAACCCTGTTCTTTTGACTCAG 3955  
Qy 448 GluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSer 467  
Db 3956 GAGCAGAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATACAGAACTCTTTG 4015  
Qy 468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487  
Db 4016 AGGAAAGCACACATCTTTGGAATTTACATTTCTTCCACCCAGGCAATGATGTATTTTCC 4075  
Qy 488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507  
Db 4076 TATGCTGGATGTTCCGTTTGGAGCTTCTTGTGGCCATTAACATCATGAGCTTTGAG 4135  
Qy 508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527  
Db 4136 GATGTTCTGTTAGTATTCTCAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4195  
Qy 528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547  
Db 4196 TCATTTGCTCTGACTATGCAAGCCAAATATATCAGCAGCCACCATCATCATGATCAT 4255  
Qy 548 GluLysLysProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGlu 567  
Db 4256 GAAAAAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4315  
Qy 568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587  
Db 4316 GGAATGTCATTTGTTGAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4375  
Qy 588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607  
Db 4376 CTTCCAGGACTGAGCTGAGGTGAAGAGGCCAGCAGCTGCTGCTGCTGCTGCTGCTGCTG 4435  
Qy 608 GlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGly 627  
Db 4436 GCGTGTGGGAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4495  
Qy 628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647  
Db 4496 AAGTGTCTGTTGATGCGCAAGAAATAAAGCGACTGAATGTTTCAAGTGGCTCCGAGCAC 4555  
Qy 648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667

Db 4556 CTGGGCATCTGTGCTCCAGGAGCCCATCTCTGTTGACGTGCAGCATTTGCTGAGAACATTTGCC 4615  
Qy 668 TxClyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAAsnAlaA 687  
Db 4616 TATGGACAAACAGCCGGTGGTGTACACGGAAGAGATCGTGAGGGCAGCAAGAGGCC 4675  
Qy 688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707  
Db 4676 AACATACATCGCTTCATCGAGTCACTGCCTTAATAAATATAGCACTAAAGTAGGACAAA 4735  
Qy 708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 727  
Db 4736 GGAACCTCAGCTCTCTGTGTGCCAGAAACAGCAATGCCATAGCTGTGCCCTTTGTAGA 4795  
Qy 728 LysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747  
Db 4796 CAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAG 4855  
Qy 748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767  
Db 4856 GTTGTCCAGAAGCCCTGGACAAAGCCAGAGAGCGCCACCTGCATTTGTGATTTGCTCAC 4915  
Qy 768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787  
Db 4916 CGCCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGAAATGGCAGAGTCAAG 4975  
Qy 788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807  
Db 4976 GAGCATGGCAGCATCAGAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGT 5035  
Qy 808 AlaGlnSer 810  
Db 5036 GTCCAGGCT 5044

RESULT 15

US-09-917-800A-1560  
; Sequence 1560, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917, 800A  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1560  
; LENGTH: 3912  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus

; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_012690  
US-09-917-800A-1560  
Alignment Scores:  
Pred. No.: 1,31e-229 Length: 3912  
Score: 2431.50 Matches: 470  
Percent Similarity: 76.03% Conservative: 155  
Best Local Similarity: 57.18% Mismatches: 184  
Query Match: 59.61% Indels: 13  
DB: 10 Gaps: 2  
US-09-873-409-2 (1-812) x US-09-917-800A-1560 (1-3912)  
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
Db 1386 ATCGATGGGCAGGACATCCGGAACCTTTAACTGTCAGGTGTCTAAGGGAATTCATCGGGCGTG 1445  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
Db 1446 GTGAGTCAAGAGCGGTACTGTTCTCTACCCAGATTGCTGAAATATCCGCTATGGCCGT 1505  
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAAspGluAlaAsnAlaValTyrAspPhe 61  
Db 1506 GGAATGTAACTATGATGAGATTAAAGAGCTGTCAAGAGGCTAATGCTATGACTTC 1565  
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81  
Db 1566 ATCATGAACTGCCACAGAAATTTGACACCTGTTGGTGACAGAGGGCGCAGCTGAGC 1625  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
Db 1626 GGGGGACAGAAACAGAGGATCGCATTTGCTGTCCTTGGTCCGCAACCCCAAGATCCTC 1685  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
Db 1686 CTGCTGACGAGGCCACGTCAGCTTGGACAGAAAGCGAAGCTAGGTGACAGGCCGCT 1745  
Qy 122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141  
Db 1746 CTGGATAAGCCAGAGAGCGCGACCACTCTGTATAGCTCACCGACTGTCAACTGTC 1805  
Qy 142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161  
Db 1806 CGGAATGCAGATGTCATCGCTGGGTTTGGAGATGGGCTCATCGTGGAGCAAGGAAGCAC 1865  
Qy 162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln----- 178  
Db 1866 AGTGAGCTGATNAAGAAAGGAGGATCTACTTCAGACTTGTTAACATGCAGACATCAGGA 1925  
Qy 179 -----AspIleLysLysAlaAspGluGlnMetGluSerMet 190  
Db 1926 AGCAGATCCTGTGACAGAGNAATTTGAAGTTGAGCTAAGTATGANAAGGCTGTGGAGGT 1985  
Qy 191 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 210  
Db 1986 GTGGCCCAAAATGGCTGGAAGCAGCGCATATTATTAGGAATTTCTACGAAGAAAGTCTGAAA 2045  
Qy 211 SerAsp-----PheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeu 228  
Db 2046 AGTTACGGGGCGCATCAAAATAGGCTGGATGGAAACCAATGAACCTTGATGANAAGCTG 2105  
Qy 229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValVal 248  
Db 2106 CCACCAAGTGTCTTTTCTGAAGGCTTAAAGCTGAATAAACAACAGATGGCCCTACTTTGTG 2165  
Qy 249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268  
Db 2166 GTGGGACACTCTGTGCCATTGCAAGCGGGCCCTCCAGCGCGCATTTCTCCATCATCTCTG 2225  
Qy 269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGlu 288  
Db 2226 TCAGATGATAGTATCTTTTGGCCCTTGGGGATGACACAGTAAGCAAGCAAGAGTGTAAAC 2285

[illegible]

Qy	649	AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr	666
Db	3366	GGCAATTGTGCCAGAGGCCATCTCTTTTGTTCGAGCATCGCCAGAAATCGCCTAC	3425
Qy	669	GlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn	688
Db	3426	GGAGACAACAGCGTGTCTCAGGATGAGATTGTGAGGCGGCCCAAGGAGGCCAAC	3485
Qy	689	IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly	708
Db	3486	ATCCACCCCTTCATTGAGACATGCCCCCAAAAGTATGAACAAGAGTAGGAGACAAGGGG	3545
Qy	709	AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLysGlnLys	728
Db	3546	ACACAGCTCTCTGGGCCAGAAACAGAGGATTGCTATCGCCCGAGCCCTCATCAGACAG	3605
Qy	729	ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal	748
Db	3606	CCTCGGCTCTACTGTGGATGAAGCCACGTCGGCTTTGGACACTGAGAGTGAAAGGTC	3665
Qy	749	ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg	768
Db	3666	GTCCAGGAAGCGTGGACAAAGCCAGGAAGGCCGACCTGCATTGTGATCGCACCGC	3725
Qy	769	LeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGlu	788
Db	3726	CTGTCCACCATCCAGAACCGCAGACTTGATCTGTGGTGATCGACACCGCAAGGTC	3785
Qy	789	GlnGlyThrHisGlnGluLeuLeuArgAsnArgPheIleTyrPheLysLeuValAsnAla	808
Db	3786	CACGGCACCCACAGCAGCTGTCTGGCCCAAGAGCATCTATTCTCCATCGTCAACATT	3845
Qy	809	GlnSer 810	
Db	3846	CAAGCT 3851	

Search completed: March 31, 2003, 16:00:18  
Job time : 444.562 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 239.871 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3323  
Sequence: 1 MIAEKGAAHMAKRGLYYS.....QELLNRNDIYKLVNAQSVQ 659

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115424\_19290/app\_query.fasta\_1.7544  
-DB=Published Applications NA -Qfmt=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09873409@cgn 1.1.960 @runat\_27032003\_115424\_19290  
-NCPU=6 -ICPU=3 -NO\_XIPXY -NO\_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:\*  
1: /cgn2\_6/ptodata/2/pubnpa/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubnpa/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubnpa/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubnpa/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubnpa/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3323	100.0	2066	10	US-09-873-409-9
2	3323	100.0	2856	10	US-09-873-409-10
3	3323	100.0	3177	10	US-09-873-409-12
4	3323	100.0	3621	10	US-09-873-409-14

Alignment Scores:  
Pred. NO.: 0  
Score: 3323.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 10

Length: 2066  
Matches: 659  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

#### ALIGNMENTS

RESULT 1

US-09-873-409-9  
; Sequence 9, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Savegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873.409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 2066  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-9

5	3323	100.0	3702	10	US-09-873-409-13	Sequence 13, Appl
6	1924.5	57.9	3924	10	US-09-880-107-2299	Sequence 2299, Ap
7	1882.5	56.7	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
8	1882	56.6	4369	10	US-09-769-097-1	Sequence 1, Appl
9	1882	56.6	4425	10	US-09-769-097-3	Sequence 3, Appl
10	1872	56.3	3860	10	US-09-866-866A-1	Sequence 1, Appl
11	1872	56.3	3860	10	US-09-866-866A-3	Sequence 3, Appl
12	1872	56.3	4643	9	US-10-072-621-2	Sequence 2, Appl
13	1872	56.3	8630	10	US-09-306-417-1	Sequence 1, Appl
14	1872	56.3	8630	10	US-09-306-417-2	Sequence 2, Appl
15	1871	56.3	4788	10	US-09-866-866A-7	Sequence 7, Appl
16	1870	56.3	4317	9	US-10-044-671-1	Sequence 1, Appl
17	1844.5	55.5	4189	10	US-09-866-866A-5	Sequence 5, Appl
18	1813	54.6	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
19	1516.5	45.6	5010	10	US-09-917-800A-483	Sequence 483, App
20	1254	37.7	4002	9	US-09-758-828-1	Sequence 1, Appl
21	1254	37.7	4002	9	US-09-758-828-3	Sequence 3, Appl
22	1251.5	37.7	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
23	1235.5	37.2	2681	10	US-09-749-340-8	Sequence 8, Appl
24	1232	37.1	4175	10	US-09-749-340-3	Sequence 3, Appl
25	1226	36.9	2698	10	US-09-749-340-5	Sequence 5, Appl
26	1211.5	36.5	3512	10	US-09-749-340-7	Sequence 7, Appl
27	1098.5	33.1	3861	9	US-09-938-842A-263	Sequence 263, App
28	1053.5	31.7	4653	9	US-10-101-388-2	Sequence 2, Appl
29	1053.5	31.7	8036	9	US-10-101-388-1	Sequence 1, Appl
30	840	25.3	3792	9	US-09-882-694-10	Sequence 10, Appl
31	836.5	25.2	1810	10	US-09-749-340-4	Sequence 4, Appl
32	828	24.9	1940	10	US-09-873-409-16	Sequence 16, Appl
33	812.5	24.5	2021	10	US-09-873-409-15	Sequence 15, Appl
34	783	23.6	3999	9	US-09-882-694-9	Sequence 9, Appl
35	758.5	22.8	2298	9	US-10-156-239-18	Sequence 18, Appl
36	758.5	22.8	2298	10	US-09-795-693-18	Sequence 18, Appl
37	758.5	22.8	3408	9	US-10-156-239-16	Sequence 16, Appl
38	758.5	22.8	3408	10	US-09-795-693-16	Sequence 16, Appl
39	758.5	22.8	3512	9	US-10-072-621-1	Sequence 1, Appl
40	710.5	21.4	1749	10	US-09-815-242-6018	Sequence 6018, Ap
41	708.5	21.3	1749	10	US-09-815-242-9908	Sequence 9908, Ap
42	693.5	20.9	1812	10	US-09-815-242-4147	Sequence 4147, Ap
43	693	20.9	1761	9	US-10-260-877-25	Sequence 25, Appl
44	693	20.9	1764	10	US-09-815-242-6863	Sequence 6863, Ap
45	649	19.5	2247	9	US-09-870-759-40	Sequence 40, Appl

US-09-873-409-1 (1-659) x US-09-873-409-9 (1-2066)

Qy 1 MetLeuAlaGluLysGlyValAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 87 ATGCTGGCGGAGAAAGGACCATGCTGAACCTAATGGCAAAACGAGGCTCTATTATTCA 146  
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40  
Db 147 CTTGTGATGTCACAGGATATTAATAAGCTGATGAACAGATGGAGTCAATGACATATTCT 206  
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60  
Db 207 ACTGAAAGAAAGACCAACTCACTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 266  
Qy 61 IleAspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80  
Db 267 ATTGCAAGGCTGAGGAATCCACCCCAATTAAGAGATAAAGTCTTCTGAACTCTCTTA 326  
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100  
Db 327 TTAAAAATTTAAAGTTAAACAGCTGATGGCTTTTGTGGTCTCTGGGACATTTGGCT 386  
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120  
Db 387 TCTGTCTTAATAGGAAGTCTCATCCAGTATTTTCCATCATCTTTGCAAAAAATTATAAC 446  
Qy 121 MetPheGlyAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140  
Db 447 ATGTTTGGAAATTAATATAAACCACTTAAGACATGATGCAGAAATTTATTCATGATA 506  
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160  
Db 507 TTCGTCAATTTGGTGTATTGCTTTGTGCTTATTTTCATGCGGGAATTTATTTACGCG 566  
Qy 161 ArgAlaGlyGluLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180  
Db 567 AGAGCAGGGGAAATTTAAACGATAGATTAAAGACATTTGGCCCTTCAAGGCCATGTTAT 626  
Qy 181 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrIleLeu 200  
Db 627 CAGGATATTGGCTGGTTTGATGAAAGGAAACAGCACAGGAGGCTTGACAAATTTA 686  
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220  
Db 687 GCCATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATGGCGCTTTACACAA 746  
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240  
Db 747 AATGCAACTTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACA 806  
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260  
Db 807 TTCCCTGATTCAGTATGCTCCAGTACTTGGCGTGACAGGAATGATGAAACCGCAGCA 866  
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280  
Db 867 ATGACTGGATTTGCCAACAAAGATAGCAAGAACTTAAGCATGCTGGAAGATAGCAACT 926  
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300  
Db 927 GAAGCTTTGGAGAAATATACGTACTATAGTGTCTATTACAAAGGGAAGCCCTTCGAGCAA 986  
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320  
Db 987 ATGTATGAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAACACAGATATT 1046  
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340  
Db 1047 GGAAGCTGTATGCAATTCAGCATGCTTTATATATTTTGGCTATCCAGCAGGGTTTCA 1106  
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360  
Db 1107 TTTGGAGCCTATTATTAATCAAGCTGGACGAATGACCCCGAGGGCATGTTTCAGTGT 1166

Qy 361 ThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380  
Db 1167 ACTGCAATTCATATGAGACTATGGCCATCGGAAAAACGCTCGTTTGGCTCCTCAATAT 1226  
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400  
Db 1227 TCCAAAGCAAAATCGGGGCTCGCATCTGTTTGGCTTGTGGTGGTGGTGGTGGTGGT 1286  
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420  
Db 1287 GACACCCAGCTCAAGAAAGGGAAGCCAGACATGATGAAGGAATTTAGAGTTTCA 1346  
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
Db 1347 GAACTCTCTTCTTCTTATCCATGTCGCCAGATGTTTTTCATCTCTCGTGGTTCCTC 1406  
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
Db 1407 AGTATTGAGCGAGGAAAGACAGTAGTATTTGTGGGAGCAGCGCTGTGGGAAAGCACT 1466  
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlyGlnValLeuPheAspGly 480  
Db 1467 TCTGTCAACTTCTGCAGAGACTTTTATGACCCCGTGAAGGACAAAGTGTGTTTATGAT 1526  
Qy 481 ValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGln 500  
Db 1527 GTGGATGCAGAAAGATTCATGATGACAGTGGCTCCGTTCCTCAATAGCAATCGTTCCT 1586  
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
Db 1587 GAGCGTGTGCTCTTCAACTGCAGCATTCGTGAGAACATCGCTATGTTGACAAACGCGT 1646  
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnIleAlaTyrGlyAspAsnSerPheIle 540  
Db 1647 GTGGTGCCATTTAGATGAGATCAAGAGCCGCAATGAGCAAAATATCCATCTCTTTAT 1706  
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGly 560  
Db 1707 GAAGGCTCTCCCTGAGAAATACACACACAGTTGGACTGMAAGGAGCACAGCTTTCGCG 1766  
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580  
Db 1767 GCCCAGAAACAAAGACTAGCTATTGCAAGGCTCTCTCCAAAAACCCAAAAATTTATTG 1826  
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600  
Db 1827 TTGGATGAGGCCACTTCAGCCCTCGATATGACATGAGAGGAGTGGTTCAGCATGCCCTT 1886  
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620  
Db 1887 GATAAGCCAGGACCGGAGGACATGCTTAGTGGTCACTCAGAGCTCTCTGCAATTCAG 1946  
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640  
Db 1947 AACGCAGATTTGATAGTGGTCTGCACAATGGAAGAAAGAAAGAAAGAAAGAAAGAA 2006  
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659  
Db 2007 GAGCTCTCGAGAAATCGAGACATATTTTAAAGTTAGTGAATGCACAGTCAGTCGAGC 2063

## RESULT 2

US-09-873-409-10  
; Sequence 10. Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05



; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-10

## Alignment Scores:

Pred. No.: 0 Length: 2856  
Score: 3323.00 Matches: 659  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-1 (1-659) x US-09-873-409-10 (1-2856)

Qy	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	877	ATGCTGGCGGAGAAAGGACACATGCTGAACCTAATGGCAAAACGAGGTCATATATTATTC	936
Qy	21	LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer	40
Db	937	CTTGTCATGTCACAGGATATTAAAGCTGATGAACAGATGGAGTCAATGACATATTCT	996
Qy	41	ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe	60
Db	997	ACTGAAAGAAAGACCACTCACTTCCTGCACTCTGGAAGAGATCAAGTCAGACTTC	1056
Qy	61	IleAspLysAlaGluLysThrGlnSerLysGluIleSerLeuProGluValSerLeu	80
Db	1057	ATTGCAAGGCTGAGGAATCCACCACCTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTA	1116
Qy	81	LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla	100
Db	1117	TAAAGATTTTAAAGTTAAACAGGCTGAATGGCGCTTTTGTGGTCTCGGGGACATTTGGCT	1176
Qy	101	SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr	120
Db	1177	TCTGTCTAAATGGAACTGTTTCATCCAGTATTTTCATCACTTTCGCAAAATTTATAC	1236
Qy	121	MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleLysSerMetIle	140
Db	1237	ATGTTTGGAAATTAATGATAAAACACATTAAGAGATGATGCAGAAATTTATTTCCATGATA	1296
Qy	141	PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly	160
Db	1297	TTCTGTCAITTTGGGTGTTATTTTGTCTTGTCTAGTTATTTTCATGAGGGATTTATTTACGGC	1356
Qy	161	ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr	180
Db	1357	AGACAGGGGAAATTTTACGATGATGATTAAGACACTTGGCCCTTCAAGCCCATGTTATAT	1416
Qy	181	GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu	200
Db	1417	CAGGATATTGCTGTTGATGAAAGGAAACACACACAGGAGGCTTGACAAATATTA	1476
Qy	201	AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln	220
Db	1477	GCCATAGATATAGCAAAATTCAGGAGCAACACAGGTTCCAGGATTTGGCGCTTTAAACAA	1536
Qy	221	AnaAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr	240
Db	1537	AATGCAACTTAACATGGGACTTTTCAGTATATCTTTCATATATGATGGAGATGACA	1596
Qy	241	PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla	260
Db	1597	TTCTCGATTCTGAGTATGCTCCAGTACTTGGCGTGACAGGAATGATGAACCCGACGA	1656
Qy	261	MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr	280
Db	1657	ATGACTGGATTTGCCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT	1716

Qy	281	GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlyLysAlaPheGluGln	300
Db	1717	GAAGCTTTGGAGAATATACGTACTATAGTGTCTATTAAACAGGAAAAAGCCTTCGAGCAA	1776
Qy	301	MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle	320
Db	1777	ATGTATGAAGAGATGCTTCAGACTCAACACAGAATACCTCGAAGAAAGACAGATTAAT	1836
Qy	321	GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg	340
Db	1837	GGAGCTGTTATGTCATTCAGCCATGCTTTATATATTTTGCCTATGACAGAGGTTTCGA	1896
Qy	341	PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe	360
Db	1897	TTTGAGGCTATTAAATTCAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTT	1956
Qy	361	ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr	380
Db	1957	ACTCAATTTGCATATGGAGCTATGCCATCGGAAAAACGCTCGTTTGGCTCCTGAATAT	2016
Qy	381	SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle	400
Db	2017	TCCAAAGCCAAATCGGGGCTGCGCACTCTGTTGCTTGTGGAAAAAGAACCAATATA	2076
Qy	401	AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg	420
Db	2077	GACAGCGCAGTCAAGAGGAAAAAGCCACACATGTGAAGGAAATTTAGAGTTTCGA	2136
Qy	421	GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu	440
Db	2137	GAAGTCTCTTTCTTATCCATGTGCCCATGTTTTCATCTCCTCCGTGGCTTATCCCTC	2196
Qy	441	SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr	460
Db	2197	AGTATTGAGCCAGGAACACAGTAGCATTTGTGGGACAGCGGCTGTGGGAAAAAGCACT	2256
Qy	461	SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly	480
Db	2257	TCTGTTCAACTTTCGACAGACTTTTATGACCCCGTGCAAGGACAAAGTGTGTTGATGT	2316
Qy	481	ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln	500
Db	2317	GTGGATGCAAAAGAAATTAATGTACAGTGGCTCGGTTCGCAAAATAGCAATCGTTCCCTCAA	2376
Qy	501	GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg	520
Db	2377	GAGCTGTGCTCTTCACTGACGACATGCTGAGAAACATCGCTATGGTGACAAACGCGT	2436
Qy	521	ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle	540
Db	2437	GTGGTGCCATTTAGATGAGATCAAGAGCCGCAAAATGACAGCAAAATATCCATTTTATT	2496
Qy	541	GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly	560
Db	2497	GAAGGTCTCCCTCGAGAAATACAAACACAAAGTTGGACTGAAAGGAGACAGCTTTCTGC	2556
Qy	561	GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu	580
Db	2557	GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCAAAACCACCAAAATTTATTG	2616
Qy	581	LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu	600
Db	2617	TTGGATGAGCCACTTCAGCCCTCGATATGACAGTGAGAGGTGTTGAGCATGCCCTT	2676
Qy	601	AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln	620
Db	2677	GATAAGCCAGGACGGGAGGACATGCTAGTGGTCACTCAAGGCTCTCTGCAATTCAG	2736
Qy	621	AnaAlaAspLeuIleValValLeuHisGlnGlyLysIleLysGluGlnGlyThrHisGln	640
Db	2737	AACGAGATTTTGATAGTGGTCTGTCACAAATGGAAAGATAAGGAAACAGGAACTCATCAA	2796

Qy	641	GlulLeuLeuATGAsnATcGAspIleTyrPheLysLeuValAsnAlaGlnSerValGln	659
Db	2797	GAGTCCTCGAGAAATCGAGACATATATTTAAAGTTAGTGAATGCACAGTCAGTCAG	2853
RESULT 3			
US-09-873-409-12			
; Sequence 12, Application US/09873409			
; Patent No. US20020037522A1			
; GENERAL INFORMATION:			
; APPLICANT: Frank, Markus			
; APPLICANT: Savegh, Mohamed			
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein			
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof			
; FILE REFERENCE: 81994/268611			
; CURRENT APPLICATION NUMBER: US/09/873,409			
; CURRENT FILING DATE: 2001-06-05			
; NUMBER OF SEQ ID NOS: 19			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 12			
; LENGTH: 3177			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: No. US20020037522A1e			
; LOCATION: (198)..(198)			
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)			
US-09-873-409-12			
Alignment Scores:			
Pred. No.:		0	Length: 3177
Score:		3323.00	Matches: 659
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		10	Gaps: 0
US-09-873-409-1 (1-659) x US-09-873-409-12 (1-3177)			
Qy	1	MetLeuAlaGluLysGlyValAlaHisAlaGluLeuMetAlaLysAspGlyLeuTyrTyrSer	20
Db	1198	ATGCTGCGGNGAAGAGACACATGCTGAATTAATGCAAAACGAGGTCTATATTATCA	1257
Qy	21	LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer	40
Db	1258	CTTGATGTGCACAGGATATTTAAAAAGCTGATCAACAGATGGAGTCAATGACATAATCT	1317
Qy	41	ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe	60
Db	1318	ACTGAAGAAGAACCACTCACTTCCTCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTC	1377
Qy	61	IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu	80
Db	1378	ATTGACAAGCTGAGGAATCCACCAATCTTAAGAGATAAGTCTTCTGGAAGTCTCTCTA	1437
Qy	81	LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla	100
Db	1438	TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTCTCTGGGACATTGGCT	1497
Qy	101	SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr	120
Db	1498	TCGTCTCTAAATGAACTGTTTCATCCAGTATTTTCCATCATCTTTTGCAAAAATTTATAACC	1557
Qy	121	MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle	140
Db	1558	ATGTTTTCGAAATAATATGATAAAACACATAAAGCATGATGCAGAAATTTATTCATGATA	1617
Qy	141	PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly	160
Db	1618	TTCGTCATTTTGGGTGTATTATTTGCTTTGTCACTTATTTTCATGCGGATATTTTACGGC	1677
Qy	161	ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr	180
Db	1678	AGAGCAGGGGAAATTTTAAACATGAGATTAAGACATTTGGCTTTCAAGGCCATGTTATAT	1737

```
Qy 541 GluGlyLeuProGluLysTyAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 2818 GAAGGTCTCCCTGAGAAATACAAACACAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 2877
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 580
Db 2878 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTATTG 2937
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 2938 TTGGATGAGGCCACTTCAGCCCTCGATTAATGACAGTGAAGAGTGGTTTCAGCATGCCCTT 2997
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 2998 GATAAAGCCAGGACGGGAAGGACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCCAG 3057
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 3058 AACGCAGATTTGTAGTGGTTCTGCACAAATGGAAGATAAAGAAACAAAGGAACCTCATCA 3117
Qy 641 GluLeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSerValGln 659
Db 3118 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAAGTGCAG 3174
RESULT 4
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human p-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14
Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-873-409-1 (1-659) x US-09-873-409-14 (1-3621)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyPyrSer 20
Db 1642 ATGCTGGCGGAGAAAGGAGCACATCTGAACCTAATGGCAAAACGAGTCTATATTATTC 1701
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyPyrSer 40
Db 1702 CTTGTGTGTCTCAGAGATATTAATAAAGCTGATGAACAGATGGAGTCAATGACATATTTCT 1761
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1762 ACTGAAAGAAAGAACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1821
Qy 61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
Db 1822 ATTGACAGGCTGAGGAATCCACCACCAATCAAGAGATGAAGTCTTCCCTGAAGTCTCTCA 1881
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValLeuGlyThrLeuAla 100
Db 1882 TTAAAAATTTTAAAGTTAAACAAGCTCGAATGGCCCTTTTGTGGTCTCTGGGGACATTGGCT 1941
```

```
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1942 TCTGTCTCTAAATGAACTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTTATAACC 2001
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyPyrSerMetIle 140
Db 2002 ATGTTTGGAAATAATGATAAAACACATTAAGCATGATGCAGAAATTTATTTCATGATA 2061
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyPyrPheMetGlnGlyLeuPheTyGly 160
Db 2062 TTCTGTATTTGGGTGTTATTTGCTTTGTCTAGTTATTTTCATGCAGGGATTTATTTCAGGC 2121
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyPyr 180
Db 2122 AGACGAGGGGAAATTTTAAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 2181
Qy 181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 2182 CAGGATATTGCTGGTTTGGTGAAGAAAGAAACAGCAGGAGGCTTGCACAACTATTA 2241
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 2242 GCCATAGATAGCACAAATTCAGAGGACACAGGTTCCAGGATTTGGCGTCTTAACACAA 2301
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyPyrGluMetThr 240
Db 2302 AATGCAACTAATCATGGGACTTTTCAGTTATCATCTTTTATATATGATGGAGAGATGACA 2361
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 2362 TTCCTGATTTCTGAGTATTTGCTCCAGTACTTCCCGTGACAGGAATGATTGAAACCGCAGCA 2421
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 2422 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAAGATAGCACT 2481
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2482 GAAGCTTTGGAGAAATATACGTACTATAGTGTCTAATTAACAAGGAAAGAGCCTTCGAGCAA 2541
Qy 301 MetTyGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2542 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCAGAGATTA 2601
Qy 321 GlySerCysTyPyrAlaPheSerHisAlaPheIleTyPyrPheAlaTyPyrAlaAlaGlyPheArg 340
Db 2602 GGAAGCTGTTATGATTCATTCAGCCATGCCCTTTATATATATTTTGGCTATGCAGCAGGGTTT 2661
Qy 341 PheGlyAlaTyPyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2662 TTTGGAGCCTATTTAAATTTCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTT 2721
Qy 361 ThrAlaIleAlaTyPyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyPyr 380
Db 2722 ACTGCAATTGCAATATGGAGCTATGCGCATCGGAAAAACGCTCGTTTTGCTCCTCGAATAT 2781
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2782 TCCAAAGCCAAATCGGGGCTGGCGCATCTGTTTGCCTTTGGTGAAGAAACCAAAATATA 2841
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2842 GACAGCCGACGTCAAGAGGGGAAAAAGCCACACACATGTGAAGGAAATTTAGAGTTTCGA 2901
Qy 421 GluValSerPhePheTyPyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2902 GAAGTCTCTTTCTATCCATGTCGCGCAGATGTTTTCATCCTCCGTGGGCTTATCCCTC 2961
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 2962 AGTATTGGAGCGGAAAGACAGTAGCATTTTGTGGGAGCAGCGGCTGTGGGAAAAAGCACT 3021
```

QY 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480  
Db 3022 TCGTTCAACTCTCGAGAGACTTTATGACCCCGTGAAGGACAAAGTGTGTTGATGGT 3081  
QY 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500  
Db 3082 GTGGATGCAAAAGAAATGAATGTACAGTGGCTCGTTCCTCAATAGCATCGTTCTCTCA 3141  
QY 501 GluProValLeuPheAsnGlnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
Db 3142 GAGCCTGTGCTTCACTGACGCAATGTGTGAGAACATCGCTTATGATGACACAGCCGT 3201  
QY 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIle 540  
Db 3202 GTGGTGCCATTAGATGAGATCAAAAGGCGCAAAATGACAGCAATATCCATTCTTTTATT 3261  
QY 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560  
Db 3262 GAAGGTCCTCCCTGAGAAATACACACACAAAGTTGGACTGAAAGGACACAGCTTCTGGC 3321  
QY 561 GlyGlnGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580  
Db 3322 GGCAGAAACAAAGACTAGCTATGCAAGGGCTCTCTCCAAAAACCCAAATTTTATTG 3381  
QY 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600  
Db 3382 TTGGATGAGGCCATTTCAGCCCTCGATATGACAGTGAAGAGTGGTTTTCAGCATGCCCTT 3441  
QY 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620  
Db 3442 GATAAGCCAGGACGAGGAGGACATGCTAGTGGTCACTCAGAGCTCTCTGCAATTTCAG 3501  
QY 621 AsnAlaAspLeuLeuValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGln 640  
Db 3502 AACCGAGATTGATAGTGGTCTGCAAAATGGAAGATGAAGGAAACAAAGCACTCATCA 3561  
QY 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659  
Db 3562 GAGCTCCTGAGAAATCGACATATATTTTAAAGTTAGTGAATGACAGTCAGTGCAG 3618

RESULT 5

US-09-873-409-13  
; Sequence 13, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 3702  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (723)..(723)  
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)  
US-09-873-409-13

Alignment Scores:  
Pred. No.: 0 Length: 3702  
Score: 3323.00 Matches: 659  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-1 (1-659) x US-09-873-409-13 (1-3702)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 1723 ATGCTGGCGGAGAAAGGACACATGCTGAACATAATGGCAAAACGAGGTCTATATTATTTCA 1782  
QY 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40  
Db 1783 CTGTGATGTACAGAGATATTAAAAAGCTGATGAACACATGGAGTCAATGACATATTCT 1842  
QY 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60  
Db 1843 ACTGAAAGAAAGAACCCACTCCTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1902  
QY 61 IleAspLysAlaGluGlnSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80  
Db 1903 ATTGCAAGGCTGAGAAATCCACCAATCTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTA 1962  
QY 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100  
Db 1963 TTAATAATTTTAAAGTTAAACCAAGCTGAATGCCCTTTTGTGGTCTTCTGGGACATTTGGCT 2022  
QY 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120  
Db 2023 TCTGTTCTAAATGGAACTGTTTCATTCAGTATTTTTCATCATCTTTTGCAGAAATATTATACC 2082  
QY 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140  
Db 2083 ATGTTTGGAAATTAATGATAAAACCAACATTAAGCATGATGACAGAAATTTTATCCATGATA 2142  
QY 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160  
Db 2143 TTTCGTCATTTTGGGTGTTTATTTGCTTTTGTGCTTATTTTTCATGACGGGATTTATTTACGGC 2202  
QY 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180  
Db 2203 AGAGCAGGGAATTTTAAACGATGAGATTAAGACACTTGGCCCTTCAAGCCATGTTATAT 2262  
QY 181 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200  
Db 2263 CAGGATATTCCTGCTGTTTGTATGATAAAGGAAACACAGCAGGAGGCTTGACAAATATTA 2322  
QY 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220  
Db 2323 GCCATAGATATAGCACAAATTAAGGAGCAACAGGTTCCAGGATTTGGGCTTTTAAACACAA 2382  
QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240  
Db 2383 AATGCAACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACA 2442  
QY 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260  
Db 2443 TTCTCTGATTCGAGTATTGCTCCAGTACTTGCCTGACAGGAATGATTGAACCCGACGA 2502  
QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280  
Db 2503 ATGACTGGATTGTCACCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCACT 2562  
QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300  
Db 2563 GAAGCTTTGGAGATATACGTACTATAGTGTCTATTAAAGGGAAGAAAGCCCTTCGAGCAA 2622  
QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320  
Db 2623 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTAT 2682  
QY 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340  
Db 2683 GGAAGCTGTATGCAATTCAGCCATGCTTTATATATTTTTCCTATGCTATGACGAGGGTTTCA 2742  
QY 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360  
Db 2743 TTTGGAGCTATTTAATTTCAAGCTGGAGCAATGACCCAGAGGGCATGTTCTATAGTTTTT 2802

Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380  
Db 2803 ACTGCAATTGCATATATGAGCTATATGCCATCGGAAACGCTGTTTTGGCTCCTGCAATAT 2862  
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400  
Db 2863 TCCCAAGCCAAATCGGGGCTCGCATCTGTTTGGCTTCTTGGTGGAAAGAACCAATATA 2922  
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420  
Db 2923 GACACCGCAGTCAGAGGGAAGGAAAGCCAGACACATGTGAAGGCAATTTAGAGTTTCGA 2982  
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
Db 2983 GAAAGTCTCTTTCTTATCCATGTGCGCCAGATGTTTTTCATCCTCCGTGGCTTATCCCTC 3042  
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
Db 3043 AGTATTGACGAGGAAGACAGTATTTTGGGGAGCAGCGCTGTGGGAAAGCACT 3102  
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 480  
Db 3103 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAGTGTGTTTGATGGT 3162  
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500  
Db 3163 GTGGATGCAAAAGAAATTGAATGACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCTCAA 3222  
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
Db 3223 GAGCCTGTGCTCTTCAACTGACGACATGCTGAGAACATCGCCTATGGTGACAAACGCCGT 3282  
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540  
Db 3283 GTGGTGCCATTAGATGAGATCAAGAGCGCGCAATGCAGCAAAATATCCATTTCTTTATT 3342  
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560  
Db 3343 GAAGGTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGCG 3402  
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580  
Db 3403 GCCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCCAAAATTTATTG 3462  
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600  
Db 3463 TTGGATGAGCCCACTTCAGCCCTCGNATATGACGTGAGAAAGTGGTTCAGCATGCCCTT 3522  
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620  
Db 3523 GATAAAGCCAGCAGCGGAAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3582  
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640  
Db 3583 AACGAGATTTGATAGTGTCTGCACAATGGAAGATAAAGGAATAAAGGAACACTCATCAA 3642  
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659  
Db 3643 GAGCTCTCTGAGAAATCGACATATATTTTAAGTAGTGAAATGCACAGTCAGTGCAG 3699

RESULT 6  
US-09-880-107-2299  
; Sequence 2299, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2299  
; LENGTH: 3924  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234  
US-09-880-107-2299

Alignment Scores:  
Pred. No.: 5,2e-195 Length: 3924  
Score: 1924.50 Matches: 372  
Percent Similarity: 74.10% Conservative: 123  
Best Local Similarity: 55.69% Mismatches: 162  
Query Match: 57.91% Indels: 11  
DB: 10 Gaps: 3

US-09-873-409-1 (1-659) x US-09-880-107-2299 (1-3924)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 1851 GTAAATTGGAGCAAGGAAGCCACAGCAACTGATGAAGAAGGAAGGGGTGTACTTCAA 1910  
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31  
Db 1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGATTTGAACATAATGAT 1970  
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51  
Db 1971 GAAAGGCTGCCACTAGAAATGCCCAAAATGCTGGAATCTCGCTATTATTAGGCATTCT 2030  
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 59  
Db 2031 ACTCAGAAAAACCTTAAATAATTCACAAATGTGTCAAGAAGAGCCTTGTATGTGGAAACCGAT 2090  
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89  
Db 2091 GGACTTGAAGCAAAATGTGCCACAGTGTCTTTCTGAAGGTCCTGAACTGAAATAAACA 2150  
Qy 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109  
Db 2151 GAATGGCCCTACTTGTGTGGGAACAGTATGTGCCATTTGCCAATGGGGGCTTCAGCG 2210  
Qy 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129  
Db 2211 GCATTTTCAGTCATATCTCAGAGATCATAGCGATTTTTTGGACCAAGCGATGATGAGTG 2270  
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149  
Db 2271 AAGCAGCAGAGTGCACATATCTTTTGATTTTCTTATTCTGGGAATATTATTTCTTTT 2330  
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169  
Db 2331 TTTACTTTCTTCCTTCAGGGTTTCAGTTTGGAAAGCTGGCGAGATCTCTCACCAGAGA 2390  
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189  
Db 2391 CTGCGGTCAATGGCTTTTAAAGCAATGTCTAAGCAGGACATGAGCTGGTGGTTGATGACCAT 2450  
Qy 190 GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209  
Db 2451 AAAAAACAGTACTGGTGCACTTTCTCAAGACTTGCACAGATGTGCCCAAGTCCCAAGGA 2510  
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229  
Db 2511 GCCACAGGACCGGTTGGCTTTAATTCACAGATATAGCTAACTTGGAACTGGTATT 2570  
Qy 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuLeuSerIleAlaProVal 249

```
Db 2571 ATCATATCATTTATACCGTGGCAGTTAAACCTATTGCTATAGCAGTTGTTCCAAATT 2630
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2631 ATTGCTGTGTCAGCAATTTGTAATGAATTTGGCTGGAAATGCCAAAGAGATAAA 2690
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAargThrIle 289
Db 2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGCGCAATAGAAAAATATTAGGACAGTT 2750
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuThrGln 309
Db 2751 GTGTCTTTGACCAGGAAAGAAATTTGAATCAATGTATGTTGAAAAATTTGATGACCT 2810
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSerHisAla 329
Db 2811 TAGAGGAATTTCTGCGAGAGGACACATCTTGGAAATTAATCTTTTAGTATCTCACAGCA 2870
Qy 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGly 349
Db 2871 TTTATGTATTTTCTATGCGGTGTTGTTTCGATTTGGTGTCATATCTCATTTGGAATGGA 2930
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db 2931 CATATCGCTTCAGAGATGTTATTTCTGGTGTCTTCTGCAATTTGATTTGGTGCGGT 2990
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db 2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTTAAGCTAAGCTGTCTGCAGCCAC 3050
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db 3051 TTATTATGCTGTTTGAAGACAACTCTGATTGACAGCTACAGTCAAGAGGGCGTGAAG 3110
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db 3111 CCTGATAAAATTTGAAGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCAACCGA 3170
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db 3171 GCNAAGTCGCGAGTCTTACGGGGCTGAGCTGGAGGTGAAGAAAGCCAGACACTAGCC 3230
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db 3231 CTGTGTGGCAGCAGTGGCTGTGGGAAGACACGGTGTCCAGCTCTCGAGCGGTCTTAC 3290
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3291 GACCCCTTGGCGGGACAGTGTCTTCGATGGTCAAGAAAGCAAAAGAACTCAATGTCCAG 3350
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db 3351 TGCTCAGAGCTCAACTCGGAATCGTGTCTCAGAGGCCATCTCTATTGACTGACGAT 3410
Qy 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db 3411 GCGGAGAATTTGCTTATGAGACACACGCGGGTGTATCACAGGATGAAATTTGTGAGT 3470
Qy 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db 3471 GCAGCCAAAGCTGCCAACATACATCTCTTTTCATCGAGACGTTACCCCAAAATATGAACA 3530
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db 3531 AGAGTGGGAGATGAAGGGACTGAGCTCTCAGGAGGTCAAAAACAGAGGATGCTATTGCC 3590
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db 3591 CGAGCCCTCATCAGACAACTCAAACTCTCTCTGTTGATGAAGCTACATCAGCTCTGGAT 3650
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
```

```
Db 3651 ACTGAAGTGAAGAGTTGTTCAGAGCCCTGGACAAAGCCAGAGAGGCCGACCTGC 3710
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db 3711 ATTGTGATTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAG 3770
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
Db 3771 AATGGAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGCACAGAAAGGCATCTAT 3830
Qy 650 PheLysLeuValAsnAlaGlnSer 657
Db 3831 TTTTCAATGGTCAGTGTCCAGGCT 3854

RESULT 7
US-09-917-800A-1560
; Sequence 1560, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1560
; LENGTH: 3912
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012690
US-09-917-800A-1560

Alignment Scores:
Pred. No.: 1,56e-190 Length: 3912
Score: 1882.50 Matches: 363
Percent Similarity: 73.43% Conservative: 129
Best Local Similarity: 54.18% Mismatches: 165
Query Match: 56.65% Indels: 13
DB: 10 Gaps: 2

US-09-873-409-1 (1-659) x US-09-917-800A-1560 (1-3912)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 1842 GTCATCGTGGAGCAAGGAAGCCACAGTGCATGATGAAGAAAGGAGGATCTACTTCAGA 1901
Qy 21 LeuValMetSerGln-----AspIleLysLys 29
```



Db 1902 CTTGTTAAATCGACATCGAAGCAGATCCTGTCAGAAAGAAATTTGAAGTTGAGCTA 1961  
Qy 30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49  
Db 1962 ACTGATGAAAGGCTGCTGGAGGTGGGCCCAATGGCTGGAAAGCAGCATATTAGG 2021  
Qy 50 LeuHisSerValLysSerIleLysSerAsp-----PheIleAspLysAlaGluSer 67  
Db 2022 AATTCTACGAAGAAAGTCTGAAAGTTTCAACGGGGCCATCAAAATAGGCTGGATGGAA 2081  
Qy 68 ThrGlnSerLysGlnLysSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
Db 2082 ACCAATGAACCTTGATGCAACAGTGCACAGGTCTTTTCTGAAGTCTTAAGACTGAAT 2141  
Qy 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
Db 2142 AAAACAGAGTGGCCCTACTTTGTGGTGGGACACCTGTGGCCATGCCAACGGGGCCCTC 2201  
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127  
Db 2202 CAGCGGCATCTCCATCATCTGTCAGAGATCATAGCTATCTTTGGCCCTGGGGATGAC 2261  
Qy 128 ThrThrLeuLysHisaspAlaGluIleLysSerMetIlePheValIleLeuGlyValIle 147  
Db 2262 ACAGTAAAGCAACAGAGGTAAACATGTTCTCGCTGGTCTTCTTGGGCCCTAGGATCCAC 2321  
Qy 148 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 167  
Db 2322 TCCTTCTTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2381  
Qy 168 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 187  
Db 2382 ACAAGGCTCCGGTCCATGCGCTTCAAGCAATGCTAAGACAGGACATAGCTGGTTGAC 2441  
Qy 188 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 207  
Db 2442 GATCATAAACACAGTACTGGTGGCTCTCTACAAGACTCGCCACAGAGCTCGGAGGTC 2501  
Qy 208 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 227  
Db 2502 CAAGGACACAGGACACAGGTTGGCTTAAATGTCACAGACACACAGCCCACTTGGAAAG 2561  
Qy 228 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerIleAla 247  
Db 2562 GGTATTATTATATATTATTATTACGGTTGGCAACTGCACACTTCTGCTTATCAGTTGT 2621  
Qy 248 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 267  
Db 2622 CCATTCTGCTGTAGCGGAAATGTTGAAATGTAATAATGTTGGTGGCAACGCCAAGAGA 2681  
Qy 268 AspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAsp 287  
Db 2682 GATATAAAGAGATGGAAGCTCTGGAAGATGTCACAGAGGCAATAGAAATATTGCG 2741  
Qy 288 ThrIleValSerLeuThrArgLysAlaPheGluGlnMetTyrGluGluMetLeuGln 307  
Db 2742 ACTGTTGTATCTTGACCAAGAGAGAAAATTTAGTCAATGTATGTTGAAAAATTACAC 2801  
Qy 308 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSer 327  
Db 2802 GGACCTTACAGGAATTCAGTGGGAAGGCTCATCATCTAGCGCATCATCTTTAGCATCTCA 2861  
Qy 328 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 347  
Db 2862 CAAGCATTCATGACTTTTCTTATGCTGGCTGCTTTCGATTTGGTTCTTACCTCATTTGG 2921  
Qy 348 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 367  
Db 2922 AATGGACATCGCTTCAAGGATGTCATCTGGTGTCTCAGCAATCGTGTGGTGCA 2981  
Qy 368 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 387  
Db 2982 GTGGCTCTAGGACATCCAGCTCATTTGTGCTCCAGACTATGCAAAAGCCAAAGCTGTCTGCA 3041

Qy 388 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 407  
Db 3042 GCATCTATTATTCAGTCTGTTTGAAGACAACTCTGATTGACAGCTACAGCAGAGAGGA 3101  
Qy 408 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 427  
Db 3102 ATGTGGCCGGATAAGTTTGAAGGAAGCTGACATTCATTAAGTTGTGTTCATATCCC 3161  
Qy 428 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 447  
Db 3162 ACCGGGCCAATGTGCGAGTGTCTTTCAGGGGCTGACCTCGAGGTGAAGAGGGGAGACC 3221  
Qy 448 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 467  
Db 3222 CTGGCCCTGTGGGAGTAGTGGCTGCGGAAGAGACACCGTGTCCAGCTGCTCGAGCGC 3281  
Qy 468 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 487  
Db 3282 TTCTACGACCCCATCGCCGGAACAGTGTCTCTCGATGTGTCAGGAAGCAAAAGAACTCAAT 3341  
Qy 488 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 507  
Db 3342 GTCCAGTGGTCCGAGCTCACTTGGCATTTGTGCCAGGAGCCCATCTCTGTTGACTGC 3401  
Qy 508 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIle 527  
Db 3402 AGCATCGCCCAAGAACATCGCTACGGAGACAAACACGCGTGTCTGCTCAGGATCAGATT 3461  
Qy 528 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 547  
Db 3462 GTGAGGGGGGCCCAAGAGGCCAACATCCACCCCTTTCATTGAGACACTGCCCCAAAAGTAT 3521  
Qy 548 AsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 567  
Db 3522 GAAACAGAGTAGGACAGAGGGGACACAGCTCTCTGAGGGCCAGAAACAGAGGATTGCT 3581  
Qy 568 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 587  
Db 3582 ATCGCCCGAGCCCTCATCAGACAGCTCGGGTCTTACTGCTGATGAAGCCACGTCGGCT 3641  
Qy 588 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 607  
Db 3642 TTGGACACTGAGAGTGAAGAGTCTGTCAGGAAGCGCTGGACAAAGCCAGGAGGCCGC 3701  
Qy 608 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 627  
Db 3702 ACTCGATTGTGATCGGCACCGCTGTCACCANTCCAGAACCGCAGACTTGATCTGGTG 3761  
Qy 628 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAsp 647  
Db 3762 ATCGACAACGGCAAGTCAAGGAGCACGGCACCCACCCAGCAGCTGCTGCCCCAGAAAGCC 3821  
Qy 648 IleTyrPheLysLeuValAsnAlaGlnSer 657  
Db 3822 ATCTATTCTTCATGCTCAACATTCAAGCT 3851

## RESULT 8

US-09-769-097-1  
; Sequence 1, Application US/09769097  
; Patent No. US20020055128A1  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Ellens  
; APPLICANT: John Anthony Feild  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: GP-50009-C2  
; CURRENT APPLICATION NUMBER: US/09/769,097  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 09/208,809



```

; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

```

Alignment Scores:			
Pred. No.:	2.1e-190	Length:	4369
Score:	1882.00	Matches:	362
Percent Similarity:	74.37%	Conservative:	157
Best Local Similarity:	53.95%	Mismatches:	158
Query Match:	56.64%	Indels:	14
DB:	10	Gaps:	5
US-09-873-409-1 (1-659) x US-09-769-097-1 (1-4369)			
QY	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTySer	20
DB	2140	GTCAATTGGGAGCAAGAAATCATGATGAGCTCATGAGAGAGAAGAAATTACTTCAAA	2199
QY	21	LeuValMetSerGln-----AspIleLysLysAlaAspGluLmMetGluSer	36
DB	2200	CTTGTCATGACTCAGACACGAGAAATGAATGAATTAGGAATGAAGCTTGTCATCT	2259
QY	37	-----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----	48
DB	2260	AAAGACGGAAATTGATATGTGGACATGCTCTTCAAAGAAGATTCGGATCCAGCTCAATAGA	2319
QY	49	ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysThr	68
DB	2320	AGAAGATCAACTCGCAAAAGCATCCGTGGCCACATGATCAAGACGGGAACCTTAGCACC	2379
QY	69	---GlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn	87
DB	2380	AAAGAGGCTCTGGATGACGACGCTACTCTCAGCTTCCTTTTGGCGGATCTCGAAGTTGAAT	2439
QY	88	LysProGluTirProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	107
DB	2440	TCAACTGAATGGCTTATTTTGGTGTGTATTTTGGCCATAATAATGGAGGCTTG	2499
QY	108	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys	127
DB	2500	CAACACGACATCTCCATAATATTTTCAAAGTGTGTAGGGGTTTTTACAAAAAATGACACC	2559
QY	128	ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	146
DB	2560	CCTGAAATCCAGCGGACAGAACCACTGTTTTCTTATTGTTCTCTGATCCCTTGGGATC	2619
QY	147	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLeuIleLeu	166
DB	2620	ATCTCTTTTCAATACGTTTTTCTTCAGGCTTCACATTTGGCAAAAGCTGGAGAGATCCTC	2679
QY	167	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	186
DB	2680	ACCAAGCAGACTCGATACATGTTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTT	2739
QY	187	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	206
DB	2740	GATGACCCCTAAAAACACACACAGAGCGCTGACCCACGAGCTTGCCAATGACGCTGCTCAA	2799
QY	207	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	226
DB	2800	GTGAAGGGGGCTACAGGGCTTAGGCTTGCCTGTATTATCCAGACATAGCAAAATCTTGGG	2859
QY	227	LeuSerValIleIleSerPheIleTyrGlyTirpGluMetThrPheLeuIleLeuSerIle	246

Db	2860	ACAGGCATCATATATCCCTGATCTACGGCTGGCAATTTGACACATTTTACTCTCTACGCAATT	2911
Qy	247	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	266
Db	2920	GTTCCTCATCTGCTATACAGGAGTGGTTGAAATGAAATGTTGTTCTGGACAAGCGGTG	2979
Qy	267	LysAspIysGlnLeuLysHisAlaGlyIysIleAlaThrGluAlaLeuGluAsnIle	286
Db	2980	AAAGATAAGAAAGAACTAGAAAGTTCTCGGAGATCGCTCAGAGAAGCAATTTGAAACATTT	3039
Qy	287	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	306
Db	3040	CGCACTGTCGTCTTTTGACTCGGAGACAGAAGTTTGAAACTATGATGATGCCACGAGCTTG	3099
Qy	307	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	326
Db	3100	CAGATACCATACAGAAATCTTTGAAGAAAGCGCAGCTTTTGGGATCACATTTCTCTCTTC	3159
Qy	327	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIle	346
Db	3160	ACCCAGGCCATGATGATTTCTCTCATGCTGCTGTTTCGSGTTTGATGCCCTACTCTGGTG	3219
Qy	347	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly	366
Db	3220	GCAGGAGAACTCATGACATTTGAAATGTTCTGTGTAGTATTCTCAGCTATTGTCTTTGGT	3279
Qy	367	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	386
Db	3280	GCCATGGCAGTGGGCGAGTCAAGTTCAATCGCTCTGACTACGCCAAGCCAAAGTCTCG	3339
Qy	387	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	406
Db	3340	GCATCCCACTCATCAGGATCATTCAGAAAAATCCCTGAGATTGACAGCTACAGCACGGAG	3399
Qy	407	GlyLysIysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	426
Db	3400	GGCTTGAAGCCTTAATATGTTGGAAGAAATGTGAAATTTAATGAGTCAATGTTCAACTAT	3459
Qy	427	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	446
Db	3460	CCACCCGACCCCAACATCCAGTGTTCAGGGCTGAGCCTTAGAGGTGAAGAAAGGCGAG	3519
Qy	447	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	466
Db	3520	ACGCTGGCGCTCTGTGGGCACAGTGGCTCGCGGAAGAGTACAGTGGTCCAGCTGCTTGAG	3579
Qy	467	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	486
Db	3580	CGCTTCTATGACCCCATGGCCGGAACAGTGTTCCTAGATGGCAAGAAATTAAGACCACTC	3639
Qy	487	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	506
Db	3640	AATGTCAGTGGCTCGCGGCCACCTGGGCATTTGTCTCCAGGAGGCCATCTCTTTTCAC	3699
Qy	507	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	526
Db	3700	TGCAGCATCCCGAGAAACATTGCCTACGGAGACACAGCCGCTGCTGTCTCATAAAGGAG	3759
Qy	527	IleIysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
Db	3760	ATCGTGAAGAGCCCAAGAGGCCCAACATCCACCAAGTTTCATCGACTCACTGCCCTGAGAAA	3819
Qy	547	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
Db	3820	TACACACACAGTGGGACAAAGGGACTCAGCTGTTCGGCGGGCAGACGACGCGATC	3879
Qy	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	586
Db	3880	GCCATCGCGCGGCCCTCTCTCAGACAGCTCATCTTACTTCTTGGATGAAGCGACATCA	3939
Qy	587	AlaLeuAspAsnAspSerGlnLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
Db	3940	GCCTCTGGATACGGAGTAGTGAAGAGTCTCCAGGAAGCGCTGGCAACAGCCAGGGAAGGC	3999

```
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4000 CGCACCTGATGTGATCGCGACCGCTGCTCCACCATCCAGACGACAGCTTGATCGTG 4059
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 646
Db 4060 GTGATTCAGAACGCGCAGGTCAAGGACGACGCGACCCACCGACGAGCTGCTGGCCCGAGAAA 4119
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 4120 GGCATCTATTTCTCGATGCTGAGTGTGAGGCT 4152
RESULT 9
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3
Alignment Scores:
Pred. No.: 2,14e-190 Length: 4425
Score: 1882.00 Matches: 362
Percent Similarity: 74.37% Conservative: 137
Best Local Similarity: 53.95% Mismatches: 158
Query Match: 56.64% Indels: 14
DB: Gaps: 5
US-09-873-409-1 (1-659) x US-09-769-097-3 (1-4425)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 2140 GTCATTGTGGAGCAAGGAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAA 2199
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer 36
Db 2200 CTTGTCTGACTCAGACAGCAGGAAATGAATTAATTAATTAATTAATTAATTAATTAAT 2259
Qy 37 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48
Db 2260 AAAGAYGGAATTGATAATGTGGCATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGA 2319
Qy 49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluSerThr 68
Db 2320 AGAAGATCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGGAACCTTAGCACC 2379
Qy 69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2380 AAAGAGGCTCTGGAGACAGCATCTCCAGCTTCTTTTGGCGGATCTCTGAAGTTGAAT 2439
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
```

```
Db 2440 TCAACTGAATGGCCTTATTTTGGTGGTGTATTTTGTGCCATAATAAATGGAGCGTTG 2499
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db 2500 CAACGACGATCTTCATATATATTTTCAAGGTTGTGAGGGTTTTTACAAAAAATGACACC 2559
Qy 128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2560 CCTGAATTCACGCGGACAGACAGCAACTTGTCTTTATTTGTTCTCTGATCCTTGGGATC 2619
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2620 ATCTCTTTTCATTACGTTTTCCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCTCTC 2679
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
Db 2680 ACCAAGCGACTCCGATACATGCTTCAATCCATGCTGAGACAGGACATTAAGCTGGTTT 2739
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2740 GATGACCTTAAAAACACACACAGGAGCGCTGACACAGGCTTGCCTCAATGACGCTGCTCAA 2799
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2800 GTGAAGGGGCTACAGGCTGCTGCTTATTTATCCCGAAGCATATAGCAAAATCTTGGG 2859
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
Db 2860 ACAGGATCATCATATCCCTGATCTACGGCTGGCAATGACACTTTTACTCTTAGCAATT 2919
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 2920 GTTCCCATCATGCTATAGCAGGAGTGGTGAATGAAATGTTGTCTGGCAAGCGCTG 2979
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2980 AAAGATAAGAGGAACATAGAAAGTTCTGGGAAGATCGCTACAGAAGCAATTTGAAAAC 3039
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 3040 CGCACTGTCTCTCTTTGACTCGGAGCAGAAAGTTTGAACATATGATGCTCCAGAGCTTG 3099
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3100 CAGATACCATACAGAAATGCTTTGAAGAAAGCGCAGCGCTTTGGGATCACATTTCTCTTC 3159
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIle 346
Db 3160 ACCCAGGCGCATGATGATTTCTCTATGCTGCTGTTTCCGGTTTGTATGCTCTACTTGGT 3219
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3220 GCACAGAACTCATGACATTTGAATAATGTTCTGTAGTATTTCTAGCTATTTGTCTTTGGT 3279
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3280 GCCATGGCAGTGGGCGAGGTTCAGTTTCCTCTCTGACTACGCGAAAGCCAAAGCTCTCG 3339
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3340 GCATCCCACTCATCAGGATCATTTGAGAAAAATCCCTTGAGATTGACAGCTACAGCAGGAG 3399
Qy 407 GlyLysLysProAspThrCysGluLysAsnLeuGluPheArgGluValSerPhePheTyr 426
Db 3400 GCGTTGAAGCCCTAATATGTTGGAAGAAATGTGAATTTAATGGAGTCTATGTTCAACTAT 3459
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3460 CCCACCCGACCAACATCCAGTCTTCAGGGGCTGACCTTAGAGGTGAAGAAAGGCGAG 3519
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
```

```
Db 3520 AGCTGGCCCTCGTGGCGACGATGGCTCGGGAGAGATACAGTGGTCCAGCTGCTTGAG 3579
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlnLeu 486
Db 3580 CGCTTCATGACCCCATGCCGGAACAGTGTTCATAGATGGCAAGAAATAAGCAACTC 3639
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3640 AATGTCAGTGGCTCGCGCCACCTCGGCAATGTGTCCAGGAGCCCATCTGTTTGAC 3699
Qy 507 CysSerIleAlaGlnLeuIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3700 TGAGCATCGCCGAGAACATGTGCTACGAGAACACAGCCGTGTCTGTCTCATAGGAG 3759
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3760 ATCGTAAGCGCAGCCCAAGGAGCCACATCCACCATCTCATCGACTCATCGCTGAGAA 3819
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3820 TACAACACAGAGTGGGAGACAAAGGAGCTCAGCTGTCCGGCGGCGAGAGCGCATC 3879
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3880 GCATCGCGCGCCCTCGTCAGACAGCCCTCATCTTACTTCTGTGATGAAGCGCATCA 3939
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3940 GCTCTGATACGAGAGTGAAAGGTCTGTCAGGAGCGCTGGACAAAGCGGGAAGGC 3999
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4000 CGCACCTGCTGATGTCGCGCACCCGCTGTCCACCATCCAGAACCGACAGACTTGATCGTG 4059
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 4060 GTGATTCAGAACGGCGAGGTCAAGGAGCGACGGCACCCACGAGCAGCTGCTGGCCAGAAA 4119
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 4120 GGCATCTATTCTCGATGTCAGTGTGCAGGCT 4152

RESULT 10
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.: 2,02e-189 Length: 3860
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
```

```
DB: 10 Gaps: 4
US-09-873-409-1 (1-659) x US-09-866-866A-1 (1-3860)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1813 GTCATTGTGGAGAAAGAAATCATGATGAATCATGAAGAGAAAGGCACTTACTTCAA 1872
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 1873 CTGTGCACAAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 1932
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 1933 AAAAGTGAATTTGATGCTTGGAAATGCTTCAAATGATTCAGATCCAGTCTTAATAAGA 1992
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 1993 AAAAGATCAACTGTAGGAGTGTCCGTGGATCAAGGCCCAAGACAGCAAGAAAGTTAGTACC 2052
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2053 AAAGAGGCTCTGATGAAAGTATATCTCCAGTTTCCTTTTGAGGATTTATGAAGCTAAAT 2112
Qy 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2113 TTAACCTGAATGGCCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2172
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAsp 126
Db 2173 CAACAGAGATTTGCAATATATTTTCAAGATTTATAGGGTTTTTACAAGATTTGATGAT 2232
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2233 CCTGAAACAAACAGACAGAAATAGTAACTTGTGTTTTCATGTTGTTGTTGTTGTTGTTG 2292
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2293 ATTCTTTTATTACATTTTCTTCAAGTTTTCATGTTGCAAGAGTGGAGAGATTCCTC 2352
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 2353 ACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGTTT 2412
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2413 GATGACCTTAAACACACCACTGGAGCATTTGACTACCAAGCTCGCCAATGATGCTCAA 2472
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2473 GTTAAAGGGGCTATAGGTTCCAGGCTTGTGTAATTTACCCAGATATAGCAATCTTGG 2532
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db 2533 ACAGAAATTAATTAATTCCTTCACTATGTTGTCGAACCTAACACTGTTACTCTTAGCAAT 2592
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 2593 GTACCCATCATTCGAATAGCAGGAGTTGTTGAAATGAAATGTTGTCGACAGCACTG 2652
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2653 AAAGATAAGAAAGAACTAGAGGTTGCTGGAAAGATGCTTACTGAAAGCAATAGAAAACCTTC 2712
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 2713 CGAACCGTTGTTTCTTTGACTCAGAGCAGAGATTTGAACATATGATGCTCAGAGTTTG 2772
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 2773 CAGGTACCATACAGAAACTCTTTGAGAAAGCACACATCTTTTGGAAATTAACATTTCTCTC 2832
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
```

```
Db 2833 ACCCAGGCAATGATATATTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCCTACTGTGTG 2892
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleValGly 366
Db 2893 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTGTCTTTGGT 2952
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 2953 GCCATGGCGGTGGGCAAGTCAGTTCAATTTGCTCTGACTGTCACAAAGGCCAAATATCA 3012
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGlyLysProAsnIleAspSerArgSerGlnGlu 406
Db 3013 GCAGCCCATCATCATCATCATCATGATTCATGAAACCCCTTTGATTGACAGCTACAGCAGGAA 3072
Qy 407 GlyLysLysProAspThrCysGluGlyAenLeuGluPheArgGluValSerPhePheTyr 426
Db 3073 GGCCTAATGCCGAAACATATGGAGGAATATGTCATTTGGTGAAGTTGTATTCACTAT 3132
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3133 CCCACCCGACGGACATCCAGCTGTCTCAGGACTGAGCTGGAGGTGAAGAAGGCCAC 3192
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 3193 ACGCTGGCTCTGTGTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCGGAG 3252
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3253 CGGTCTTACAGCCCTTGGCAGGAAGTGTCTGTATGGCAAGAATAAAGCGACTG 3312
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3313 AATGTTAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTGTTTGAC 3372
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3373 TGCAGCATTTGCTGAGAACATTCCTATGTGAGACAAAGCCGGGTGTGTACAGGAAGAG 3432
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3433 ATCGTGAGGCGAGCAAGAGAGCCCAACATACATGCTTCATCGATGCTACTGCTTAATAA 3492
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3493 ATAGCACTAAGTAGGACACAAAGAACTCAGCTCTCTGTGGTGGCCAGAAACAACGCA 3552
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3553 GCCATAGCTCGTGCCCTGTATAGACAGCCCTCATATTTTGTGATGAGCCACGTCA 3612
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3613 GCTCTGGATACAGAAAGTGAAGGTGTCCAAGAGCCCTGGCAAAAGCCAGAGAGGC 3672
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 3673 CGCACTGCATTGTGATGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 3732
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 3733 GTGTTTCAGAAATGGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGTGGCAGAGAA 3792
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3793 GGCATCTATTTTCAATGTGTGCTGCTCAGGT 3825
```

## RESULT 11

```
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US2002010224A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
```

```
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866.866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3
```

## Alignment Scores:

Pred. No.:	2.02e-189	Length:	3860
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	10	Gaps:	4

US-09-873-409-1 (1-659) x US-09-866-866A-3 (1-3860)

```
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1813 GTCATTGGGAGAAAGAAATCATGATCAATCATGAAAGAGAAAGGCAATTTACTTCAA 1872
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 1873 CTTGTCAATGTCAGACAGCAGGAAATGAAGTTGAATAGAAAATGCAGCTGATGAATC 1932
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 1933 AAAAGTGAATGTGCTGCTGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGA 1992
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 1993 AAAAGTCAACTCGTAGGAGTGTCCGTGATCAACAAGCCCAAGCAGAAAGCTTAGTACC 2052
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2053 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCTTTGGAGGATTATGAAGCTAAT 2112
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2113 TTAACTGAATGGCCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2172
Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2173 CAACAGCATTTTGCATAATATTTTCAAGATATATAGGGGTTTTTACAGAAATGATGAT 2232
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2233 CTTGAAACAAACGACAGATAGTAACTTGTTCCTACTATTGTTTCTAGCCCTTGAAT 2292
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2293 ATTTCTTTTATTAATTTTCTTCAAGGTTTTCATATTGGCAAGCTGGAGAGATCTCTC 2352
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 2353 ACCAAGCGGCTCCGATACATGTTTTCGATCCATGCTCAGACAGAGTGTGAGTTGGTTT 2412
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2413 GATGACCTTAAACACCACTGGAGCATTTGACTACAGGCTCGCCAATGATGCTGCTCAA 2472
```

QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
Db 2473 GTTAAAGGGGCTATAGTTTCCAGCTTGTGTAATACCAGAAATAGCAATCTTGGG 2532  
QY 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246  
Db 2533 ACAGGAATTAATATATCTTCATCTATGGTTGGCAACTAACACTGTCTTACCTTAGCAATT 2592  
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
Db 2593 GTACCCATCATTCATAGCAGAGTGTGTGAATGAATAATGTGTGACAGACACTG 2652  
QY 267 LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
Db 2653 AAGAATAAGAAAGAACTAGAAAGTGTGGAGATCGCTACTGAAGCAATAGAAAACTTC 2712  
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
Db 2713 CGAACCGTGTCTTGTGACTCAGGAGCAGAGTGTGAACATATGATGCTCAGAGTTTG 2772  
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
Db 2773 CAGGTACCATACAGAAACTCTTTGAGAAAGCACACATCTTTGGAAATTACATTTTCCTC 2832  
QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346  
Db 2833 ACCAGGCAATGATGATTTTCTATATGCTGGATGTTTCGGTGTGAGCCTACTTGGTG 2892  
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
Db 2893 GCACATAAATCTAGTGTGAGGATGTTCTGTGTAGTATTTTCACTGTGTGCTTTGGT 2952  
QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
Db 2953 GCATGGCGTGGGCAAGTCAGTTCTATTGCTCTCTGACTATGCCAAGCCAAATATCA 3012  
QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
Db 3013 GCAGCCCATCATCATGATCATTTGAAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTG 3072  
QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426  
Db 3073 GGCTTAATGCGCAACACATTGGAAGAAATGTCACATTTGGTGAAGTGTATTTCACACTAT 3132  
QY 427 ProCysArgProAspValPheIleLeuArgGlyLysSerLysSerIleGluArgGlyLys 446  
Db 3133 CCACCCGACCGACATCCAGTGTCTCAGGACTGAGCTGAGGTGGAAGAGGCCAG 3192  
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
Db 3193 ACCTGGCTCTGTGGGACAGCTGCTGTGGAGAGACACAGTGTCTCAGCTCTCTGGAG 3252  
QY 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
Db 3253 CGGTCTACAGCCCTTGGCAGGAAAGTCTCTGTATGGCAAGAAATAAAGCCACTG 3312  
QY 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
Db 3313 AATGTTCACTGGCTCCGAGCACACCTGGGACATGCTGTCCAGAGGCCATCTCTGTTGAC 3372  
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526  
Db 3373 TGCAGCATTTGCTGAGAACATTTGCTATGGAGACACAGCCGGTGTGTACAGGAAGAG 3432  
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
Db 3433 ATCGTAGGGCACAAGAGGCGCACATACATGCTTTCAGTGCCTTTCAGTGCCTAATAAA 3492  
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
Db 3493 TATAGCACTAAAGTAGGACAAAGGAACTCAGCTCTCTGTTGGCCAGAAACACCGCAT 3552  
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586

Db 3553 GCATAGCTCGTGGCCCTTGTGTAGACAGCCTCATATTTCTTTGGATGAAGCCAGCTCA 3612  
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
Db 3613 GCTCTGGATACAGAAGTGAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGGC 3672  
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
Db 3673 CGCACTGCTATTGTGCTTCCACCGCCTGTCCACCATCCAGATGCAGACTTAATAGTG 3732  
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646  
Db 3733 GTGTTTCAAGATGGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGCTGTGCGCACAGAA 3792  
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657  
Db 3793 GGCATCTATTTTCAATGTGTGCTCAGTGTCCAGGCT 3825  
RESULT 12  
US-10-072-621-2  
; Sequence 2, Application US/10072621  
; Patent No. US20020169137A1  
; GENERAL INFORMATION:  
; APPLICANT: Reiner, Peter B.  
; APPLICANT: Connop, Bruce P.  
; APPLICANT: Pollard, Michelle  
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION  
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY  
; FILE REFERENCE: 100103.402  
; CURRENT APPLICATION NUMBER: US/10/072,621  
; CURRENT FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 4643  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-072-621-2  
Alignment Scores:  
Pred. No.: 2,68e-189 Length: 4643  
Score: 1872.00 Matches: 356  
Percent Similarity: 74.22% Conservative: 142  
Best Local Similarity: 53.06% Mismatches: 159  
Query Match: 56.33% Indels: 14  
Db: 9 Gaps: 4  
US-09-873-409-1 (1-659) x US-10-072-621-2 (1-4643)  
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 2234 GTCATGTGGAGAAAGAAATCATGATGAACTCATGAAAGAGAGGAGGCAATTTACTTCAA 2293  
QY 21 LeuValMetSerGln-----AspIleLysAlaAspGluGlnMet 34  
Db 2294 CTTGTCACAATGCAGACAGCAGAAATGAAGTTGAATTAGAAAAATGCAGCTGATGAATCC 2353  
QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
Db 2354 AAAAGTGAATTTGATGGCTTGGAAATGCTTCAATATGATTCAAGATCCAGTCTAATAAGA 2413  
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67  
Db 2414 AAAAGTCAACTCGTAGGAGTGTCCGTGGATCAACAGCCCAAGAGAGGCTTAGTACC 2473  
QY 68 ThrGlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
Db 2474 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTTATGAAGCTAAAT 2533  
QY 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
Db 2534 TTAAGTGAATGGCCCTTATTTTGTGTGTATTTGTGTCATTAATAATGGAGCCCTG 2593

Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126  
Db 2594 CAACAGCATTTTGAATAATATTTTCAAGATATATAGGGTGTTCACAGAAATTCATGAT 2653  
Qy 127 LysThrLeuLysHisAspAlaGluLysSerMetIlePheValIleLeuGlyVal 146  
Db 2654 COTGAACAAACAGCAGAAATAGTAACTTGTTCACATATGTTTCTACCTCTGGAAT 2713  
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGluLysLeu 166  
Db 2714 ATTTCTTTATTAACATTTTCTTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCTC 2773  
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186  
Db 2774 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTT 2833  
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrIleLeuAlaIleAspIleAlaGln 206  
Db 2834 GATGACCCCTAAACACACCATGGAGCATTTGACTACGAGCTCGCAATGATGCTGCTCA 2893  
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
Db 2894 GTTAAAGGGCTATAGTTCAGGCTTGTGTAATATACCAGAAATATACCAATCTTGG 2953  
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246  
Db 2954 ACAGAAATATATATCTTCATCTATGTTGGCAACTTAACACTGTACTCTTAGCAATT 3013  
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
Db 3014 GPACCCATTCATTCAGGAGTGTGTAATGAAATGTTGTCTGGCAAGCACTG 3073  
Qy 267 LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
Db 3074 AAGATAGAAAGAACTAGAGGCTGCGGAAGATGCTACTGGAAGCAATAGAAAACCTTC 3133  
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeu 306  
Db 3134 CGAACCGTGTTCCTTGACTCAGGAGCAGAGTTTGAACATATGTATGCTCAGAGTTG 3193  
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
Db 3194 CAGGTACCATACAGAACTCTTTGAGAAAGCAGACATCTTTGGAAATTACATTTTCCTTC 3253  
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346  
Db 3254 ACCCAGGCAATGATGATTTTCTCTATGCTGGATGTTCCGGTTTGGAGCCTACTTGGTG 3313  
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
Db 3314 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGCTTGT 3373  
Qy 367 AlaMetAlaIleGlyThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
Db 3374 GCCATGGCGTGGGCAAGTCACTTATTTGCTCTGCTGACTATGTCACCAAGCCAAATATCA 3433  
Qy 387 AlaAlaHisLeuPheAlaLeuLysLysProAsnIleAspSerArgSerGlnGlu 406  
Db 3434 GCAGCCCATCATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTACAGACGGAA 3493  
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuPheArgLysValSerPheTyr 426  
Db 3494 GGCCTAATCCGAACACATTTGAAGAAATGTACATTTGGTGAAGTTGATTTCAACTAT 3553  
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
Db 3554 CCCACCCGACCGACATCCAGTGTCTTCAGGAGCTGAGCTGGAGTGAAGAGGCCAG 3613  
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
Db 3614 ACGTGGCTCTGGTGGCAGCAGTGGCTGTGGGAAGAGACAGTGTGCTCAGCTCTGGAG 3673

Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
Db 3674 CGTTCCTACGACCCCTTGGCAGGAAAGTGTCTGTATGGCAAGAAATAAAGCACTG 3733  
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
Db 3734 AATGTTCTAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGAGCCCATCTCTGTTGAC 3793  
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526  
Db 3794 TGCACATTCGTGAGAACATTCCTATGGAGACAACACCGCGGTGTGTGCACAGGAAG 3853  
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
Db 3854 ATTTGTAGGGCGCAAGAGGAGGCCAACATACATGCTCTTCATCGAGTCACTGCTTAATA 3913  
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeu 566  
Db 3914 TATACACTAAAGTAGGAGACAAGAACTCAGCTCTCTGTGGCCAGAAACACAGCAT 3973  
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586  
Db 3974 GCCATAGCTCGTGCCTTTAGACAGCTCATATTTTGTCTTTGATGAACCCAGTCA 4033  
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
Db 4034 GCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAAAGCCTCGCAAGCCAGAGAGGC 4093  
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
Db 4094 CGCAGCTCATTTGATTCCTCACCCTGTCACCATCCAGAAATGCAGACTTAATAGTG 4153  
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646  
Db 4154 GGTTCAGATGGCAGAGTCAAGAGATGGCAGCATCAGCAGCTGCTGCGCAGAGAA 4213  
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657  
Db 4214 GGCATCTATTTTCAATGTCAGTGTCCAGGCT 4246  
RESULT 13  
US-09-306-417-1  
; Sequence 1, Application US/09306417  
; Patent No. US20020103144A1  
; GENERAL INFORMATION:  
; APPLICANT: Heinrich-Pette-Institut  
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors  
; FILE REFERENCE: P50491  
; CURRENT APPLICATION NUMBER: US/09/306,417  
; CURRENT FILING DATE: 1999-05-06  
; EARLIER APPLICATION NUMBER: DE 198 22 115  
; EARLIER FILING DATE: 1998-05-08  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 8630  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: proviral  
; OTHER INFORMATION: plasmid DNA  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(160)  
; OTHER INFORMATION: plasmid backbone (pUC)  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (161)..(677)  
; OTHER INFORMATION: 5'-LTR  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (532)..(1219)  
; FEATURE:







Qy	587	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
		::	
Db	4832	GCTCGATACAGAAAGTGAAGGTTGTCGAAGAGCCCTGGACAAAGCCAGAGAAGGC	4891
Qy	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
		::	
Db	4892	CGCACCTGCATTTGTAATGCTACCGCCCTGTCACCATCCAGATGCACAGCTTAATAGTG	4951
Qy	627	ValLeuHisAsnGluLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	646
		::	
Db	4952	GTGTTTCAGAAATGGCAGAGTCGAAGGAGCATGCGACGATCAGCAGCTCTGGCACAGAA	5011
Qy	647	AspIleTyrPheLysLeuValAsnAlaGlnSer	657
		::	
Db	5012	GGCATCTATTTTCAATGGTCACTGTCACAGGT	5044

## RESULT 14

```

US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SPbeta91msA1

```

Alignment Scores:	
pred. No.:	6.95e-189
Length:	8630
Matches:	356
Conservative:	142
Mismatches:	159
Indels:	14
Percent Similarity:	74.22%
Best Local Similarity:	53.06%
Query Match:	56.33%

DB: 10 Gaps: 4

US-09-873-409-1 (1-659) x US-09-306-417-2 (1-8630)

QY 1 MetLeuAlaGluLysGlyValahisalaGluLeuMetAlaLysArgGlyLeuTyrSer 20  
Dbs : : : : :  
Db 3032 GTCAATTGTGGAGGAAGAATCATGATGACTCATGAAGAAGAAGGCATTTACTTCAAA 3091  
QY 21 LeuValMetSerGln-----AspIleTysLysAlaAspGluGlnMet 34  
Dbs : : : : :  
Db 3092 CTTGTACAATGCACAGCACAGCAAATGAAGTTGAATTAGAAAAATGCAGCTGATGAATCC 3151  
QY 35 GluSerMetThrTyrSerThrGluArgLysThrAenSerLeuProLeuHisSerVal--- 53  
Dbs : : : : :  
Db 3152 AAAAGTCAAAATTGATGCCVTTGGAAATGTCTTCAAAATGATTAACAAGATCCAGTCTAAATAAGA 3211  
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67  
Dbs : : : : :  
Db 3212 AAAAGATCAACTGTAGGAGTGTCCGTGGATCAGCCCCAAGACAGAAAGCTTAGTGATCC 3271  
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
Dbs : : : : :  
Db 3272 AAAGAGGCTCGATGAAGATATACCTCCAGTTCCTTTTGGAGGATTAAGAAGCTAAAT 3331  
QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
Dbs : : : : :  
Db 3332 TTAACCTCAATGGCCCTATTTCCTGTGTGTATTTTTGTGCCATATAAAATGGAGGCCT 3391  
QY 108 HisProValPheSerIlePheAlaLysIleThrMetPheGlyAsn---AsnAsp 126  
Dbs : : : : :  
Db 3392 CAACCAGCATTTGCAATAATATTTCAAAGATATATAGGGGTTTTTACAAGAAATGATGAT 3451  
QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146  
Dbs : : : : :  
Db 3452 CCTGAACAAACACGACAGATAGTAACCTGTGTTTCACATATGTTTCTAGCCCTTGGATT 3511  
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166  
Dbs : : : : :  
Db 3512 ATTTCTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 3571  
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186  
Dbs : : : : :  
Db 3572 ACCAAGCGCTCGATACATGTTTTCCGATCCATGCTCAGACAGSAGTGTAGTGTGTTT 3631  
QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThriLeLeuAlaIleAspIleAlaGln 206  
Dbs : : : : :  
Db 3632 GATGACCTTAAACACCACTGGACANTTGACTACCGGCTCGCCAATGATGCTGCTCAA 3691  
QY 207 IlecInGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
Dbs : : : : :  
Db 3692 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTAATCCAGAAATATAGCAAAATCTTGGG 3751  
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246  
Dbs : : : : :  
Db 3752 ACAGGAATAATATATPCCTTCATCTATGGTTGGCAACTAACACCTGTGTACTCTTTAGCAATT 3811  
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
Dbs : : : : :  
Db 3812 GTACCATCATTCATACAGAGAGTTGTTGNAATGAANAATGTTGTCTGGACAGCACTG 3871  
QY 267 LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
Dbs : : : : :  
Db 3872 AAGATAAGAAAGAACTAGAGGCTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3931  
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
Dbs : : : : :  
Db 3932 CGAACCCGTGTTTCTTTTGACTCAGAGACAGAGTTTGAACATATGTATGTCCTCAGAGTTTG 3991  
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
Dbs : : : : :  
Db 3992 CAGGTACCATACAGAAACCTCTTTGGAGAAAGACACACATCTTTTGAATATACATTTCTCTC 4051  
QY 327 SerHisAlaPheIleTyrPheAlaTrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346

```
Db 4052 ACCAGCAATGATGATTTTCTATGCTGGAGTGTTCGGTGTGGAGCCTACTTGGTG 4111
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleValGly 366
Db 4112 GCACATAAACTCATGAGCTTGGAGATGTTCTGTAGTATTTTCAGCTGTGTCTTGGT 4171
QY 367 AlaMetAlaIleGlyThrLeuValLeuAlaProGluTyrSerIysAlaIysSerGly 386
Db 4172 GCCATGCCCTGGGCAAGTCAGTTTCAATTTGCTTCTGACTATGCCAAGCCAAATATCA 4231
QY 387 AlaAlaHisLeuPheAlaLeuLeuGluIysPheProAsnIleAspSerArgSerGlnGlu 406
Db 4232 GCAGCCACATCATCATGATCATTTGAAAAAACCCCTTGTATTACACGCTACACACGGAA 4291
QY 407 GlyIysLysProAspThrCysGluGlyAsnLeuPheArgGluValSerPhePheTyr 426
Db 4292 GGCTAATGCCGACACATTTGGAGGAAATGTACATTTGGTGAAGTTGTATTCAACTAT 4351
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 4352 CCCACCCGACCGGACATCCCAAGTCTTCCAGGACTGAGCCTGGAGGTGAAGAAGGCCAG 4411
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyIysSerThrSerValGlnLeuLeuGln 466
Db 4412 ACCTGGCTCTGTGGGACAGTGGCTGGGAAGAGCACAGTGGTCCAGCTCCCTGGAG 4471
QY 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaIysGluLeu 486
Db 4472 CGGTTCTACAGCCCTTGGCAGGAAAGTCTGCTGTATGGTGAAGAAATAAGCGACTG 4531
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 4532 AATGTTCACTGGCTCCGAGCACACCTGGGATCGTGTCCAGGAGCCATCTCTGTTGAC 4591
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 4592 TGACGATTCCTCAGAACATTTGCTTATGGAGACAAACAGCCGGTGTGTCCAGGAAGAG 4651
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 4652 ATCGTGGGCGACCAAGGAGGCGCAACATACATGCTTCTATCAGTCACTGCCCTAATAA 4711
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 4712 TATAGCACTAAAGTAGAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACACGCATT 4771
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 4772 GCATAGCTCGTGCCTTGTATGACAGCCCTCATATTTTGTGATGAAGCCACGCTCA 4831
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 4832 GCTCTGATACAGAAAGTGAAGAGTTGTCCAGAGCCCTGCAGAACGCCAGAGAGGC 4891
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 4892 CGCACCTGCAATTTGATGTGCTCACCCGCTGTCCACCATCCAGAAATGAGACTTAATAGT 4951
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 646
Db 4952 GTGTTTCAGAAATGGCAGAGTCAAGAGGACATGGCCAGCGCATCAGCAGCTGTGGCAGAAA 5011
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 5012 GGCACTATTATTTTCAATGGTCACTGTCCAGGCT 5044
```

## RESULT 15

US-09-866-866A-7  
; Sequence 7, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian

APPLICANT: Schuetz, John  
TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
FILE REFERENCE: 1340-1-021CIP2  
CURRENT APPLICATION NUMBER: US/09/866,866A  
CURRENT FILING DATE: 2001-08-30  
PRIOR APPLICATION NUMBER: 09/584,586  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: PCT/US99/11825  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: 60/086,988  
PRIOR FILING DATE: 1998-05-28  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 7  
LENGTH: 4788  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-866-866A-7

Alignment Scores:  
Pred. No.: 3.6e-189 Length: 4788  
Score: 1871.00 Matches: 352  
Percent Similarity: 73.70% Conservative: 144  
Best Local Similarity: 52.30% Mismatches: 159  
Query Match: 56.30% Indels: 18  
DB: 10 Gaps: 4

US-09-873-409-1 (1-659) x US-09-866-866A-7 (1-4788)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 1801 GTCAATGTGCGACCAAGAAATCATGATGAGCTCATGAGAGAAAGGGCATTTACTTCAA 1860  
QY 21 LeuValMetSerGln-----AspIleLysLysAla 30  
Db 1861 CTGTCTCATCAGACAGACAGCAGCAAGAAATGAATTAATAGGAAATGAAGCTTGAATCT 1920  
QY 31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeu 50  
Db 1921 AAGGATGAAATTAATTTAGACATGCTTCAAAAGATTCCAGGATCCAGTCTAATAAGA 1980  
QY 51 HisSerValLysSerIleLysSerAspPheIle-----AspLysAlaGlu 65  
Db 1981 AGA-----AGATCAACTCGCAAAAGACATCTGTGGACCATGACCAACAGCAGAGCTT 2034  
QY 66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysLysIleLys 85  
Db 2035 AGTACCAAGAGCCCTGGATGAAGATGATCTCCAGCTTCTTTGGCGGATCTTGAAG 2094  
QY 86 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 105  
Db 2095 TTGAATTCAACTGAATGGCTTATTTTGTGGTGTATATCTGTGCCATAATAAATGGA 2154  
QY 106 ThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn--- 124  
Db 2155 GCGTTCACGCCAGCATCTCCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGGT 2214  
QY 125 AsnAspLysThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 144  
Db 2215 GGCCCCCTGAAACCCAGCGGACAGACAGCAACTGTGTTTCTGTGTGTTCTGATCCTT 2274  
QY 145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164  
Db 2275 GGGATCATTTCTTCAATATTTTCTTCCAGGGCTTCCATTTTGGCAAGCTGGAGAG 2334  
QY 165 IleLeuThrMetArgLeuArgHisIleAlaPheLysAlaMetLeuTyrGlnAspIleAla 184  
Db 2335 ATCCCTCACCAAGCGACTCCGATACATGTTTCAATATCCATCTCAGACAGGATGTGAGC 2394  
QY 185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIle 204  
Db 2395 TGGTTTGATGACCCCTAAAAAACACCACCGGAGCAGCTGACCACCGAGCTCCGCCAACGATGCT 2454

```
Qy 205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224
Db 2455 GCTCAAGTAAAGGGGTACAGGGTCTAGGCTGTCTGTGATTTTCCAGAACATAGCAAT 2514
Qy 225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244
Db 2515 CTTGGGACAGGAATCATCATATCCCTATCTATGCTGGCTGGCACTTAACACTTTTACTCTTA 2574
Qy 245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264
Db 2575 GCAATTGTACCATTCATTCGATAGCTGGAGTGTGAAATGAAATGTTGCTGTCGACAA 2634
Qy 265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284
Db 2635 GCACTGAAGATAAGAACTAGAGGTTCTCGAAGATTGTCTACGGAAGCAATTGAA 2694
Qy 285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304
Db 2695 AACTTCGCACTGTGTCTCTTTGACTCGGAGCAGAAAGTTTGAAACCATATGTATGCCAG 2754
Qy 305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324
Db 2755 AGCTTGAGATACCATACAGAAATCGCATGAGAAAGCACAGTGTGTTGGATCACGTTT 2814
Qy 325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344
Db 2815 TCCTTCAACCCAGGCCATGATGATTTTCTTATGCTGTGCTGTTTCCGGTTCGGTGCCTAC 2874
Qy 345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364
Db 2875 TTGGTGACACAACTCATGACTTTTGAAATGTTCTGTAGTATTCTCAGCTATTGTC 2934
Qy 365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384
Db 2935 TTTGGTGCATGGCAGTGGGCGAGTCAGTTCATTCGCTCTCTGACTATGCCAAGCAACA 2994
Qy 385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 404
Db 2995 GTGTGAGCATCCACATCATCAGGATCATTTAGAAATCCCGGAGATTGACAGCTACAGC 3054
Qy 405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424
Db 3055 ACGCAAGGCCCTAAAGCCGAATATGTTGGAAGGAAATGTGCAATTTAGTGGAGTCTGTGTC 3114
Qy 425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444
Db 3115 AACTATCCACCCGACCCAGCATCCAGTCTTACGGGCTGAGCCTTGAGGTGAAGAAG 3174
Qy 445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464
Db 3175 GGCCAGACGCTGGCCCTGTGTGGCAGCAGTGGCTGGGGAAGAGCACAGTGTCTCCAGCTG 3234
Qy 465 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 484
Db 3235 CTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTCTTAGATGGCAAAAGATAAAG 3294
Qy 485 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504
Db 3295 CAACGTGAATGTCCAGTGGCTCCGAGCACAGCTGGGCATTGTGTCCAAAGGCCCATTTCTC 3354
Qy 505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524
Db 3355 TTTGACTGCAGCATTCGAGAGAACATTGCTACGGAGACAAACAGCGGGTCTGTCTTAT 3414
Qy 525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544
Db 3415 GAGGAGATTGTGAGGCGACCCAAAGAGGCGCAACATCCACAGTTCATTCGACTCGCTACCT 3474
Qy 545 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGln 564
Db 3475 GATAAATAACACACGAGTAGGAGACAAAGGCATCTCAGCTCTCGGGTGGGAGAGCAG 3534
Qy 565 ArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAla 584
```

```
Db 3535 CGCATCCCATCGCACGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCA 3594
Qy 585 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
Db 3595 ACATCAGCTCTGGATACAGAAAGTGAAGGTTGTCAGGAAGCGCTGGACAAAGCCAGG 3654
Qy 605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
Db 3655 GAAGGCGCACCTGCATTTGTCGCTGCTCCGCGCTGCCACCATCCAGAACGCGGACTTG 3714
Qy 625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArg 644
Db 3715 ATCTGTGTGATTTCAGAACGGCAAGTCAAGAGCACGGCACCCACCCAGCAGCTGCTGGCG 3774
Qy 645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3775 CAGAAGGGCATCTACTTCTCAATGGTCAGTGTGCAGGCT 3813
```

Search completed: March 31, 2003, 15:57:49  
Job time : 327.871 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 8908.97 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-1  
Perfect score: 3323  
Sequence: 1 MLAEGGAHMAKGLGYY.....QELLNRDIYKLVNAQSVQ 659

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09873409/runat\_27032003.115418.19228/app\_query.fasta\_1.7544  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409@cgn.1\_54887@runat.27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*

- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rdd:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3323	100.0	2066	6	AX339027 Sequence
2	3323	100.0	2856	6	AX339028 Sequence
3	3323	100.0	3177	6	AX339030 Sequence
4	3323	100.0	3621	6	AX339032 Sequence
5	3323	100.0	3702	6	AX339031 Sequence
6	3163	95.2	3699	6	AX478104 Sequence
7	1924.5	57.9	3924	6	AX024455 Sequence
8	1924.5	57.9	3924	6	AX409652 Sequence
9	1924.5	57.9	3924	9	HUMMDR3
10	1916	57.7	3987	10	CRUPGPI185
11	1916	57.7	4296	10	CRUPGPI
12	1916	57.7	4304	10	CRUPGPI165
13	1915.5	57.6	4084	10	MUSMDRA
14	1911	57.5	4018	5	GGA9799
15	1893	57.0	3858	4	OAU78609
16	1892	56.9	4927	10	AF097746
17	1884	56.7	2752	5	AF099732
18	1882.5	56.7	3912	10	RATPLVCO
19	1882.5	56.7	3912	10	AF286167
20	1879	56.5	4323	10	AF286167
21	1876.5	56.5	3905	10	CRUPGPIII
22	1873	56.4	4186	6	AX108654
23	1873	56.4	4195	6	AX108656
24	1872	56.3	2726	6	I33621
25	1872	56.3	3840	6	AX481416
26	1872	56.3	3860	6	AX322787
27	1872	56.3	3860	6	AX322789
28	1872	56.3	3988	6	AX024454
29	1872	56.3	4192	9	AF016535
30	1872	56.3	4264	6	AR051647
31	1872	56.3	4264	6	AR051650
32	1872	56.3	4646	6	AX336420
33	1872	56.3	4646	6	AX336708
34	1872	56.3	4646	6	AX336708
35	1872	56.3	4646	6	AX391099
36	1872	56.3	4646	6	I49610
37	1872	56.3	4669	9	HUMMDR1
38	1872	56.3	4669	6	I08557
39	1872	56.3	6505	6	AR028671
40	1872	56.3	8630	6	AX012320
41	1872	56.3	8630	6	AX012321
42	1871	56.3	9318	6	AR028672
43	1871	56.3	3682	10	MUSPGLYIA
44	1871	56.3	4788	6	AX322793
45	1870	56.3	4924	10	MUSMDR1A
			4279	6	AX105057

ALIGNMENTS

AX339027 AX339027 2066 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 9 from Patent WO0194400  
DEFINITION AX339027  
ACCESSION AX339027  
VERSION AX339027.1 GI:18129119  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Frank, M.H. and Sayegh, M.H.  
AUTHORS A gene encoding a multidrug resistance human p-glycoprotein  
TITLE homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 9 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES Location/Qualifiers  
source 1..2066  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 647 a 403 c 470 g 546 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.8e-259 Length: 2066  
Score: 3323.00 Matches: 659  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-1 (1-659) x AX339027 (1-2066)  
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluMetAlaLysArgGlyLeuTyrTyrSer 20  
DB 87 ATGCTGGCGGAGAAAGGACACATGCTGAACCTAATGGCAAAACAGGCTATATATCA 146  
QY 21 LeuValMetSerGlnAspIleLysIleValAspGluGlnMetGluSerMetThrTyrSer 40  
DB 147 CTTGTGATGTACAGAGATATTAAGAGCTGATGAACAGATGGAGTCAATGACATATCT 206  
QY 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60  
DB 207 ACTGAAAGAAAGACCACTCACTCTCTGCACTCTGTAAGACATCAAGTCAAGCTTC 266  
QY 61 IleAspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80  
DB 267 ATTGACAAGGCTGAGGAATCCACCAATCTAAGAGATAGTCTTCTGAGTCTCTTA 326  
QY 81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100  
DB 327 TTAAGAAATTTTAAAGTTTAAACAGCCTGAATGGCTTTTGTGGTCTGCGGACATGGCT 386  
QY 101 SerValLeuAsnGlyThrValHisProValPheSerIleLeuPheAlaLysIleThr 120  
DB 387 TCTGTTCTAAATGAGTCTGTCAGTATTTTCCATCATCTTTGCAAAATTTATAACC 446  
QY 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140  
DB 447 ATGTTTGGAAATATATGATAAAACACATTAAGCATGATGCAGAAATTTATTCATGATA 506  
QY 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160  
DB 507 TTCGTCAATTTGGGTGTAATTTGCTTTGTGAGTATTTTCATGAGGATATTTTACGGC 566  
QY 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180  
DB 567 AGAGCAGGGGAATTTAAGCATGAGATTAAAGACACTTGGCCCTCAAGCCATGTTATAT 626  
QY 181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrIleLeu 200  
DB 627 CAGGATATTGCTGTTGTATGAAAGGAAAAACAGACAGAGGCTTGACAAATATTA 686

QY 201 AlaIleAspIleAlaGlnIleGlnGlyValaThrGlySerArgIleGlyValLeuThrGln 220  
DB 687 GCCATAGATATAGCACAAATTCAGAGGACACAGGTTCCAGANTGGCTCTTTACACAA 746  
QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240  
DB 747 AATGCAACTAACATGGGACTTTCAGTTATCATTTCTTTATATATGATGGAGATGACA 806  
QY 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260  
DB 807 TTCCTGATTTCTGAGTATTTGCTTCCAGTACTTTCCTGTCAGGAATGATTAACACGCA 866  
QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280  
DB 867 ATGACTGGATTTGCCACAAAGATAACCAAGAACTTAAGCATGCTGGAAGATAGCAACT 926  
QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300  
DB 927 GAAGCTTTGGAGAATAACGTACTATAGTGTCTAATCAAGGGAAGAAAGCTTCGAGCAA 986  
QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320  
DB 987 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTAT 1046  
QY 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340  
DB 1047 GGAAGCTGTTATGCATTCAGCCATGCTTTATATATTTTCCCTATGCGAGGTTTCGA 1106  
QY 341 PheGlyAlaTyrIleLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360  
DB 1107 TTTGGAGCTTATTTAATTTCAAGCTGGAGCAATGACCCAGAGGCGCATGTTTCATAGTTT 1166  
QY 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380  
DB 1167 ACTGCAATTCATAGAGCTATGGCCATCGGAAACCGCTCGTTTGGCTCTCGTAATAT 1226  
QY 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLysLysProAsnIle 400  
DB 1227 TCCAAAGCCAAATCGGGGCTGCGCATCTGTTTGGCTCTGTTGGAAGAAACCAATATA 1286  
QY 401 AspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420  
DB 1287 GACAGCGCAGTCAAGAGGGAAGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCA 1346  
QY 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
DB 1347 GAAGTCTCTTTCTTCTATCATGTCGCCAGATGTTTTCATCTCCGCTGCTTATCCCTC 1406  
QY 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
DB 1407 AGTATTGAGCGAGGAAAGACAGTAGCATTTTGGGGAGCGAGCGCTGTGGAAAGACACT 1466  
QY 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 480  
DB 1467 TCTGTTTCAACTCTTCGACAGACTTTATGACCCCGTCAAGGACCAAGTGTGTTGATGGT 1526  
QY 481 ValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGln 500  
DB 1527 GTGGATGCCAAAGAAATTTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAA 1586  
QY 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
DB 1587 GAGCCTGTCTCTTCAACTGCAGCATTTGTCAGAACATCGCTATGTTGTCACACACCGCT 1646  
QY 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540  
DB 1647 GTGGTGCCATTTAGATGAGATCAAGAGCCGCAATGTCAGCAAAATATTCATTTCTTTAT 1706  
QY 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGly 560  
DB 1707 GAAGGTCTCTCTGAGAAATACACACACAGTTGGACTGAAAGGAGCACAGCTTCTGCG 1766  
QY 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580

```
Db 1767 GGCACAGAACAAAGACTAGCTATTGCAAGGGCTCTTCCAAAAACCCAAAAATTTATTG 1826
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 1827 TTGGATGAGGCACCTTCAGCCCTCGATATGACATGACATGAGAGGTGGTTCAGCATGCCCTT 1886
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaLeuGln 620
Db 1887 GATAAAGCCAGGACGGAAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 1946
Qy 621 AsnAlaAspLeuValValLeuHisAsnGlyLysValLysGluGlnGlyThrHisGln 640
Db 1947 AACGCAGATTTGATAGTGGTCTGCACAAATGGAAGATAAAGGAACAAGGAACCTCATCAA 2006
Qy 641 GluLeuLeuAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 2007 GAGCTCCTGAGGAATCGACATATATTTTAAGTTAGTGAATGCACATGCTAGTGCAG 2063

RESULT 2
AX339028 2856 bp DNA linear PAT 09-JAN-2002
LOCUS AX339028
DEFINITION Sequence 10 from Patent WO0194400.
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank,M.H. and Sayegh,M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN

Alignment Scores:
Pred. No.: 2,77e-259 Length: 2856
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-1 (1-659) x AX339028 (1-2856)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 877 ATGCTGGCGGAGAAAGGACACATGCTGAACATAATGGCAAAACGAGGTCTATATTATTCA 936
Qy 21 LeuValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 937 CTTGTGATGTCAGGATATTAATAAAGCTGTATGAACACAGATGGAGTCAATGACATATTCT 996
Qy 41 ThrGluATGlyThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 997 ACTGAAAGAAAGACCACTCACTCTCTGCACTCTGTGAAGCATCAAGTCAGACTTC 1056
Qy 61 IleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80
Db 1057 ATTGCAAGGCTGAGGAATCCCAATCCCAATCTAAAGATAAGTCTTCTCCTCAAGTCTCTCA 1116
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
Db 1117 TTAATAAATTTTAAGTTAAACAAGCCGTAATGGCCCTTTTGTGGTTCGGGGACATTTGGCT 1176
```

```
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1177 TCTGTCTTAAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTTATAACC 1236
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 1237 ATGTTTGGAAATATATGATAAAACACATTAAGCATGATGACAGAAATTTATTCATGATA 1296
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 1297 TTCGTCAITTTGGGTGTTTATTTGCTTGTCACTTATTTTCATGCAGGATTTATTTACGGC 1356
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db 1357 AGAGCAGGGGAAATTTTAAACGATGAGATTAAGACACTTTGGCCCTTCAAGACCATGTTAT 1416
Qy 181 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db 1417 CAGGATATTGCTGCTGTTTGTATGAAGAAACACACAGCAGGAGGCTTGACAAATATTA 1476
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 1477 GCCATAGATATAGCACAAATTTCAAGGACCAACAGGTTCCAGGATTTGGCTCTTAACACAA 1536
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240
Db 1537 AATGCAACTAACTGGGACTTTTCAGTTATCATTTCTCTTTATATATGATGGAGATGACA 1596
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 1597 TTCTTGATTTCTGAGTATTGCTCCAGTACTTCCGTCGACAGGAATGATTTGAAACCCGACGA 1656
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db 1657 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACT 1716
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 1717 GAAAGCTTTGGAGAAATATACGTACTATAGTGTCTATTAAACAAGGGAAAAAGCCCTTCGAGCAA 1776
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 1777 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAACACAGATTTAT 1836
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 1837 GGAAGCTGTTATGCAATTCAGCATGCCCTTTATATATTTTGGCTATGCAGCAGGGTTTGA 1896
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 1897 TTTGGAGCCTATTTAATTCAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTT 1956
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 1957 ACTGCAATTCATATATGAGACTATGCCATCGAAAAACGCTGTTTGGCTCCTGAATAT 2016
Qy 381 SerLysAlaLysSerGlyValAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2017 TCCAAGGCCAAATCGGGGCTGCGCATCTGTTTGGCTTGTGGAAAAAGNAACCAATATA 2076
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2077 GACAGCCGAGCTCAAGAAGGGGAAAAAGCCAGACACATGTGAAGGGAAATTTAGAGTTTCA 2136
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2137 GAAGTCTCTTTCTTATCCATGTCGCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2196
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db 2197 AGTATTGAGCGAGGAAAGACAGTACATTTTGTGGGAGCAGCGGCTGTGGGAAAGCACT 2256
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGly 480
```



```
Db 2257 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTCGACAGGACAAGTGTCTGTTGATGGT 2316
Qy 481 ValAspAlaLysGluLeuValGlnTTPLeuArgSerGlnIleAlaIleValProGln 500
Db 2317 GTGGATGCMAAAGAAATGAATGACAGTGCTCCGTTCCCAATAGCAATGCTTCTCAA 2376
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 2377 GAGCCTGTGCTCTTCAACTGCACCAITTCGTGAGAACATCGCCTATGTTGACACACCGGT 2436
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 2437 GTGGTGCCTATTAGATGAGATCAAGAAGCGCAATATGCGCAAAATATCCATTCTTTATT 2496
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 2497 GAAGGTCTCCCTGAGAAATACACACACAGTTGGACTGAAAGAGACACAGCTTCTGGC 2556
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu 590
Db 2557 GCCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTTATTG 2616
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 2617 TTGGATGAGGCCACTTTCAGCCCTCGATAATGACAGTGAGAGGTGGTTTCAGCATGCCCTT 2676
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 2677 GATAAAGCCAGGACGGAGGACATCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 2736
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
Db 2737 AACGCAGATTGATAGTGGTCTGCAACATGGAAAGATAAAGGAACAAGGAACCTCATCA 2796
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 2797 GAGCTCTGAGAAATCGAGACATATATTTAAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 3
LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 12 from Patent WO0194400.
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
AUTHORS Frank, M. H. and Savegh, M. H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
Source 1..3177
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,2e-259 Length: 3177
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-1 (1-659) x AX339030 (1-3177)
```

```
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1198 ATGCTGGCGGAGAAAGGAGCACAATGCTGAACCTAAATGGCAAAACGAGGTCTATATTATTC 1257
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1258 CTTGTGATGTCTACAGATATTTAAAGAGCTGATGAACAGATGGAGTCAATGACATATTC 1317
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1318 ACTGAAGAAAGAACCACTCACTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1377
Qy 61 IleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80
Db 1378 ATTGAAGGCTGAGAAATCCACCAATCTAAAGAGATAAGTCTTCTCGAAGCTCTCTCA 1437
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
Db 1438 TTAAGAAATTTTAAAGTTAAACCAAGCTGAATGGCTTTTGTGGTCTTGGGACATTTGGCT 1497
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 1498 TCTGTCTAAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTGCCAAAATATAAAC 1557
Qy 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 1558 ATGTTTGGAAATTAATGATAAAACCAATTAAGAGCATGATGCAGAAATTTATTCATGATA 1617
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 1618 TTTCGTCATTTTCGGGTGTTATTTGCTTTGTCAGTTATTTTCATGAGGATTTATTACGGC 1677
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db 1678 AGAGCAGGGGAAATTTTAAAGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 1737
Qy 181 GlnAspIleAlaTTPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeu 200
Db 1738 CAGGATATTCCTGCTGTTGATGAAAGGAAACACAGCAGGAGGCTTGACAAACAATATTA 1797
Qy 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db 1798 GCCATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGGCTCTTAACACAA 1857
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 240
Db 1858 AATGCAACTAACATGGGACTTTCAGTTATTCATTTCTTATATATGATGGAGGATGACA 1917
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db 1918 TTCTCTGATTTCTGATTTGCTCCAGTACTTGCCTGACAGGAATGATTGAAACCGCAGCA 1977
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLysHisAlaGlyLysIleAlaThr 280
Db 1978 ATGACTGGATTTGGCAACAAAGATAAGCAAGAACTTAAGCACTGCTGGAAAGATAGCACT 2037
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2038 GAAGCTTTGGAGAAATATAGTACTATAGTGTCTATTAAACAGGAGGAAAGCCCTTCGAGCA 2097
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2098 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATT 2157
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 2158 GGAAGCTGTATGCAATTCAGCCATGCCCTTATATATTTTGGCTATGACAGCGGTTTCCA 2217
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2218 TTTGGAGCCTATTATTAATTCAGCTGACGATGACCCAGAGGGCATGTTTCATGTTTTT 2277
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
```

Db	2278	ACTGCAATTCATATGAGAGCTATGCCATCGAAAAACGCTGCTTTGGCTCCCTGAATAT	2337
Qy	381	SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle	400
Db	2338	TCCAAGCCAAATCGGGGCTGGCATCTGTTGGCTTTGGAAAAAGAACCAATATA	2397
Qy	401	AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg	420
Db	2398	GACAGCCGAGTCAGAAAGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCCA	2457
Qy	421	GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu	440
Db	2458	GAAGTCTCTTCTTATCTCATGTCGCCAGATGTTTTCATCTCCGCGGCTATCCCTC	2517
Qy	441	SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr	460
Db	2518	AGTATTGACGAGGAAAGACAGTACATTTTGTGGGAGCAGCGGTGTGGGAAAGCACT	2577
Qy	461	SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly	480
Db	2578	TCTGTTCAACTTCTCAGAGACTTTATGACCCCGTGAAGGACAAAGTGTGTTGATGGT	2637
Qy	481	ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln	500
Db	2638	GTGGATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCTCA	2697
Qy	501	GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg	520
Db	2698	GAGCTGTGCTCTTCACTGACGCAATGCTGAGAACATCGCTATGGTGACAAACAGCGT	2757
Qy	521	ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle	540
Db	2758	GTGGTGCCATTAGATGAGATCAAGAGCCGCAATGCGAGCAATATCCATTCTTTTATT	2817
Qy	541	GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly	560
Db	2818	GAAGTCTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGC	2877
Qy	561	GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu	580
Db	2878	GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAATTTTATG	2937
Qy	581	LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu	600
Db	2938	TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTCAGCATGCCCTT	2997
Qy	601	AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln	620
Db	2998	GATAAAGCCAGACGCGGAAGGACATGCCCTAGTGGTCACTCACAGGCTCTCTGCAATTCA	3057
Qy	621	AsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLysGluGlnGlyThrHisGln	640
Db	3058	AACGAGATTTGATAGTGGTCTCGACAAATGGAAGATGAAGAAACAGGAACACTCATCA	3117
Qy	641	GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaAsnAlaGlnSerValGln	659
Db	3118	GAGTCTCTGAGAAATCGACATATATTTTAAAGTTAGTGAATGCACAGTCACTGCAG	3174
RESULT 4			
AX339032			
LOCUS			
Sequence 14 from Patent WO0194400.			
DEFINITION			
AX339032			
ACCESSION			
VERSION			
AX339032.1 GI:18129124			
KEYWORDS			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
AUTHORS			
Frank, M.H. and Sayegh, M.H.			
TITLE			
A gene encoding a multidrug resistance human p-glycoprotein			

homologue on chromosome 7p15-21 and uses thereof			
Patent: WO 0194400-A 14 13-DEC-2001;			
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)			
FEATURES			
Location/Qualifiers			
source			
1..3621			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
BASE COUNT			
ORIGIN			
1081 a 696 c 842 g 1002 t			
Alignment Scores:			
Pred. No.: 3,81e-259 Length: 3621			
Score: 3323.00 Matches: 659			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 100.00% Indels: 0			
DB: 6 Gaps: 0			
US-09-873-409-1 (1-659) x AX339032 (1-3621)			
Qy	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	1642	ATGCTGGCGGAGAAAGGAGCACATGCTGAACCTAATGGCAAAACGAGGCTCTATATTATCA	1701
Qy	21	LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer	40
Db	1702	CTTGTGATGTCACAGGATATTAATAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCT	1761
Qy	41	ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe	60
Db	1762	ACTGAAGAAAGAACCACTCACTTCCTCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTC	1821
Qy	61	IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu	80
Db	1822	ATTGCAAGGCTGAGGAATCCACCCCAATCTAAAGAGATAAGTCTTCTGGAAGTCTCTCTA	1881
Qy	81	LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla	100
Db	1882	TTAAAAATTTTAAAGTTAAACAAGCCTGAATGCGCTTTTGTGGTTCTCGGGACATTTGCT	1941
Qy	101	SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr	120
Db	1942	TCTGTCTTAAATGGAACCTGTTCCAGTATTTTCCATCACTTTTGCAGAAATTTATAACC	2001
Qy	121	MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle	140
Db	2002	ATGTTTGAATATATGATAAAACCCACATTAAAGCATGATGCAGAAATTTATTCCATGATA	2061
Qy	141	PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly	160
Db	2062	TTCTGTCATTTTGGGTGTTATTTTGTCTAGTTATTTTTCATGTCAGGGGATTTATTTACGC	2121
Qy	161	ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheAlaMetLeuTyr	180
Db	2122	AGAGCAGGGGAAATTTTAAACCATGAGATTAAAGACACTTGGCCCTTCAAAGCCATGTTAT	2181
Qy	181	GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu	200
Db	2182	CAGGATATTGCTGGTTTGTATGAAAGGAAACACAGCAGGAGGCTTGACAACATATTA	2241
Qy	201	AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln	220
Db	2242	GCCATAGATATAGCACAAATTCAGAGGACCAACAGGTTCCAGGATTGGCGTCTTAACACAA	2301
Qy	221	AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr	240
Db	2302	AATGCAACTAACTGGGACTTTTCATGTTATTCCTTTATATATATGAGTGGAGATGACA	2361
Qy	241	PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla	260
Db	2362	TTCTGATTCCTGAGTATTGCTCCAGTACTTCCGCGTGCAGAGGATGATTGAAACCCGAGCA	2421
Qy	261	MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr	280

```
Db 2422 ATGACTGGATTTCCTCCAAAGATTAAGCAAACTTAAGCATCTCGAAAGATAGCAACT 2481
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db 2482 GAAGCTTTGAGAAATATACGTACTATAGTGTCTATTAACAGGGAAGAAAGCTTCGAGCAA 2541
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
Db 2542 ATGTATGAAGAGATGCTTTCAGATCTCAACAGAAATACCTCGAAGAAAGCACAGATTATT 2601
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 2602 GGAAGCTGTATTGCAATTCAGCCATGCCCTTTATATATTTTCCCTATGCAGAGGGTTTCGA 2661
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 2662 TTTGGAGCCTATTATTAAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTT 2721
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 2722 ACTGCAATTCATATGAGCTATGGCCATCGGAAACGCTCGTTTTTGGCTCCTGAATAT 2781
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 2782 TCCAAAGCCAAATTCGGGGGCTGGCGATCTGTTGGCCTTGTGGAAAGAAACCAATATA 2841
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db 2842 GACAGCGCAGTCAGNAGGGAAGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2901
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db 2902 GAAGTCTCTTCTTCTATCCATGTCGCCAGATGTTTCATCTCCGTGCTTATCCCTC 2961
Qy 441 SerIleGluArgGlyLysThrValAlaPheValCysSerGlyCysGlyLysSerThr 460
Db 2962 AGTATTGAGCGAGGAAGACAGTAGCATTTTGTGGGAGACGCGCTGTGGAAAGACACT 3021
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db 3022 TCTGTTCAACTTCTGCAGACACTTTATGACCCCGTGCAAGACAAAGTCTGTTGATGGT 3081
Qy 481 ValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGln 500
Db 3082 GTGGATGCAAAAGAAATTAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCTCAA 3141
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db 3142 GAGCCTGTGCTCTTCAACTCGACATTTGCTGAGAACATCGCCCTATGGTGACACAGCCGT 3201
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
Db 3202 GTGGTGCCATTAGATGAGATCAAGAGCGCAATATGCAGCAATATCCATCTTTTATT 3261
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
Db 3262 GAAGGTCTCCCTCAGAAATACACACACAAAGTTGGACTGAAAGAGCACAGCTTCTGGC 3321
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
Db 3322 GGCACAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAGCCCAAAATTTTATTG 3381
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
Db 3382 TTGATGAGGCCACTTTCAGCCCTCGATATACAGTGAAGAGTGGTTTCAGCATGCCCTT 3441
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
Db 3442 GATAAAGCCAGGACGGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3501
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyValIleLysGluGlnGlyThrHisGln 640
```

```
Db 3502 AACGCAGATTTGATAGTGGTCTTGACATGGAAGATAAAGAACAAAGAACTCATCAA 3561
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
Db 3562 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618

RESULT 5
AX339031 AX339031 3702 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 13 from Patent WO0194400.
DEFINITION AX339031
ACCESSION AX339031
VERSION AX339031.1 GI:18129123
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1..3702
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,92e-259 Length: 3702
Score: 3323.00 Matches: 659
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 6 Gaps: 0

US-09-873-409-1 (1-659) x AX339031 (1-3702)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1723 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTTAATGGCAAAACGAGTCTATATTATTC 1782
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db 1783 CTTGTGATCTCAGAGATATTAAAAAGCTGATGAACAGATGAGTCAATGACATATTCT 1842
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db 1843 ACTGAAAGAAAGAACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1902
Qy 61 IleAspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
Db 1903 ATTGCAAGCTGAGAGATCCACCNAATCTAAAGAGATGAAGTCTTCTGAGTCTCTCTA 1962
Qy 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100
Db 1963 TTAATAATTTTAAAGTTAAACAAGCTGAATGCCCTTTTGTGTTCTGGGACATTTGGCT 2022
Qy 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
Db 2023 TCTGTTTAAATGGAACTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAAAATTTATAAC 2082
Qy 121 MetPheGlyAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
Db 2083 ATGTTTGGAAATAATGATAAACCAACATTAAAGCATGATGACAGAAATTTATTCCATGATA 2142
Qy 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
Db 2143 TTCGTCAATTTGGGTGTTATTGCTTTGTCAGTTATTTTCATGCAGGATTTATTTACGCG 2202
Qy 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
```

Db 2203 AGAGCAGGGGAAATTTAAACGATGAGATTAGACACTTGGCCCTTCAAAAGCCATGTTATAT 2262  
Qy 181 GlnAspIleAlaTrrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200  
Db 2263 CAGGATATTGCTGTGTTGATGAAAGGAAACACACAGAGGGCTTGCAACAATATTA 2322  
Qy 201 AlalleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220  
Db 2323 GCCATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTTACACAA 2382  
Qy 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrrPheMetThr 240  
Db 2383 AATGCAACTAACATGGGACTTTTCAGTTATCATTTCTCTTATATATGATGGAGATGACA 2442  
Qy 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260  
Db 2443 TTCCTGATTCGAGTATTGCTCCAGTACTTGCCTGACAGGNAATGATTGMAAACCCGACGA 2502  
Qy 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280  
Db 2503 ATGACTGGATTGGCCAAACAAGATTAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACT 2562  
Qy 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300  
Db 2563 GAAGCTTTGGAGAAATATACGTACTATAGTGTCTATTAAACAGGGGAAAAAGCCCTTCGAGCAA 2622  
Qy 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320  
Db 2623 ATGTATGAGAGATGCTTCAGACTCAACACAGNAATACCTCGAAGAAAGCACAGATTATT 2682  
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340  
Db 2683 GGAAGCTGTATGCAATTCAGCCACTGCTTTATATATTTTGCCTATGACAGAGGGTTTCTGA 2742  
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360  
Db 2743 TTTGAGGCTATTATTAATCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTT 2802  
Qy 361 ThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380  
Db 2803 ACTGCAATTCGATATGGAGCTATGGCCATCGGAAAGAGCTCGTTTGGCTCTCGAATAT 2862  
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400  
Db 2863 TCCAAAGCCAAATCGGGGCTCGCATCTGTTTGTCTGTTGAAAGAAACAAATATA 2922  
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420  
Db 2923 GACAGCCGCGAGTCAAGAAGGGGAAAAAGCCAGACACATGTGTAAGGGGAATTTAGAGTTTCA 2982  
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
Db 2983 GAAGTCTCTTTCTATCCATGTCGCCAGATGTTTTCACTCCGGGCTTATCCCTC 3042  
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
Db 3043 AGTATTGAGCGAGGAAGACAGTAGCATTTTGTGGGAGCAGCGGTGTGGGAAAGCACT 3102  
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480  
Db 3103 TCTGTTCAACTTCTCAGAGACTTTATGACCCCGTGCAGAGCAAGTGTGTTGATGGT 3162  
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrrPheArgSerGlnIleAlaIleValProGln 500  
Db 3163 GTGGATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTCTCTCAA 3222  
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
Db 3223 GAGCCTGTGCTCTTCACTGACAGCATTGTGTAGAACATCGCTTATGTTGTCACACAGCGT 3282  
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540

Db 3283 GTGGTGCCATTAGATGAGATCAAAAGAGCCCAAAATGCAGCAAAATATCCATTCTTTATT 3342  
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560  
Db 3343 GAAGGTCTCCCTGAGAAATACACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3402  
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580  
Db 3403 GGCCAGAAACAAAGACTAGCTATTTCGAGGGCTCTTCTCCAAAAACCCCAAAATTTATTG 3462  
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600  
Db 3463 TTGATGAGGCCACTTCAGCCCTCGATTAATGACAGTGAAGAGTGGTTCAGATGCCCTT 3522  
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620  
Db 3523 GATTAAGCCAGCAGCGGAAGACATGCTAGTGTCTCACTCACAGGCTCTCTGCAATTCA 3582  
Qy 621 AsnAlaAspLeuIleValValLeuHisGlnGlyLysIleLysGluGlnGlyThrHisGln 640  
Db 3583 AACGCAGATTTCATAGTGGTTCTGCACAATGGAAGATAAAAGGAACAAAGGAACATCAA 3642  
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659  
Db 3643 GAGCTCTCGAANAATCGAGACATATATTTAAGTTAGTGAATGCACAGTCACTAGTCAG 3699  
RESULT 6  
AX478104 3699 bp DNA linear PAT 12-AUG-2002  
LOCUS AX478104  
DEFINITION Sequence 34 from Patent WO0240541.  
ACCESSION AX478104  
VERSION AX478104.1 GI:22217064  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1  
AUTHORS  
Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y.,  
Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,  
Sanjanwala, M., Ramkumar, J., Arvizu C., Gietzen, K.J., Lal, P.G.,  
Azimzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,  
Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,  
Policky, J.L. and Kearney, L.  
TITLE  
Transporters and ion channels  
JOURNAL  
Patent: WO 0240541-A 34 23-MAY-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1. .3699  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7472030CB1"  
BASE COUNT 1116 a 707 c 860 g 1016 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,44e-246 Length: 3699  
Score: 3163.00 Matches: 632  
Percent Similarity: 96.05% Conservative: 1  
Best Local Similarity: 95.90% Mismatches: 0  
Query Match: 95.19% Indels: 26  
DB: 6 Gaps: 1

US-09-873-409-1 (1-659) x AX478104 (1-3699)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 1798 ATGTGGCGGAGAAAGGCGCATGCTGAATAATGCGAAAAACGAGGCTATATATTATCA 1857  
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40  
Db 1858 CTTGTGATGTACAGGATATTTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATCT 1917

QY 41 ThrGluArgGlyThrAsnSerLeuProLeuHisSerValLysSerIleLysSerPhe 60  
DB 1918 ACTGAAAGAAAGACCAACTCACTTCTGCACTCTGTGAAGACATCAAGTCAGACTTC 1977  
QY 61 IleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeu 80  
DB 1978 ATTGACAAAGCTGAGGAATCCACCCCAATCTAAAGAGATAGTCTTCTGAAGTCTCTCTA 2037  
QY 81 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 100  
DB 2038 TTTAAATTTTAAAGTTAAACAGCCCTGAATGCCCTTTGTGGTCTGGGACATTTGGCT 2097  
QY 101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleThr 120  
DB 2098 TCTGTTCTAAATGGAACCTGTTATTCATCCAGTATTTTCCATCATCTTTGCAAAAATTAAC 2157  
QY 121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleLysSerMetIle 140  
DB 2158 ATGTTTGGAAATTAATGATAAAACCAATTAAAGCATGATGCAGAAATTTATTCCATGATA 2217  
QY 141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160  
DB 2218 TTGCTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTTCATG----- 2259  
QY 161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180  
DB 2259 ----- 2259  
QY 181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200  
DB 2260 CAGGATATTGCTGCTGTTGATGAAAGAGAAACAGCACAGGAGGCTTTGACAAATATTA 2319  
QY 201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220  
DB 2320 GCCATAGATATACACAAATTCAAGGACACACAGTTCCAGGATGTCGCTCTTAACACAA 2379  
QY 221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240  
DB 2380 AATGCAACTTAACATGGGACTTTTCAGTTATCATTTTCCTTTATATATGATGGAGATGACA 2439  
QY 241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260  
DB 2440 TTCCTGATTCGAGTATGCTCCAGTACTTTCGCTGACAGGAATGATTGAACCCGACGA 2499  
QY 261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280  
DB 2500 ATGACTGGATTTCACCAACAAAGATAGCAAGAACTTAAGCATCTCGGAAGATAGCACT 2559  
QY 281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300  
DB 2560 GAAGCTTTGGAGAATATACGTACTATAGTGTCTATTAACCAAGGGGAAAGCCTTCGAGCAA 2619  
QY 301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320  
DB 2620 ATGTATGAAGAGATGCTTTCAGATCTCAACACAGAAATACCTCGAAGAAAGCACAGATTAT 2679  
QY 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340  
DB 2680 GGAAGCTGTATGATTCAGCATTCAGCCATGCTTTATATATTTTTCCTATGCGGAGGTTTCGA 2739  
QY 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360  
DB 2740 TTTGGAGCCTATTAAATTCAGCTGGACGAATGACCCAGAGGCGCATGTTTCATAGTTTTT 2799  
QY 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380  
DB 2800 ACTGCAATTCATATGGAGCATGAGGCATCGGAGAACCGCTCTGTTTGGCTCTCGAATAT 2859  
QY 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400  
DB 2860 TCCAAAGCCAAATCGGGGGCTGGGCATCTGTTGCTTTGCTTTGTTGAAAGAAACCAATATA 2919  
QY 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420

DB 2920 GACACCCGAGTCAAGAAGGGGAAAAAGCCAGACATGTGAAGGAAATTTAGAGTTTCA 2979  
QY 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
DB 2980 GAAGTCTCTTCTTCTTATCCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTC 3039  
QY 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
DB 3040 AGTATTGAGCGAGGAAGACAGTAGCATTTTGTGGGAGCAGCGCTGTGGGAAAGCACT 3099  
QY 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480  
DB 3100 TCTGTTCAACTTCTCCAGAGACTTTATGACCCCGTGAAGCAAGTGTCTGTTGATGT 3159  
QY 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500  
DB 3160 GTGGATGCMAAAGAAATTTGATGTACAGTGGCTCCGTTCCCAATAGCAATCGTCTCTCAA 3219  
QY 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
DB 3220 GAGCCTGTGCTCTTCAACTGCAGCATTTGCTGAGAACATCGCCTATGTTGACACAGCCGT 3279  
QY 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540  
DB 3280 GTGGTGCCATTAGATGAGATCAAGAAAGCCCAATGACGAAATATCCATTTCTTTAT 3339  
QY 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGly 560  
DB 3340 GAAGGTCTCCCTGAGAAATACACACACAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3399  
QY 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580  
DB 3400 GGCCAGAAACAAAGACTAGTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTATTG 3459  
QY 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600  
DB 3460 TTGSGATAGGCGCACCTTCAGCCCTCGATATGACAGTGAAGAGTGGTTCAGCATGCCCT 3519  
QY 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620  
DB 3520 GATAAAGCCAGACCGGAGGACATGCTAGTGGTCTCAGCTCAGAGCTCTCTGCAATTCAG 3579  
QY 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640  
DB 3580 AACGCGAGATTTGATAGTGTCTGACAAATGCAAAAGATAAAGGAAACAGGAACTCATCAA 3639  
QY 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659  
DB 3640 GAGCTCTGAGAAATCGAGACATATTTTAAAGTTAGTGAATGACAGTCAGTGCAG 3696

## RESULT 7

LOCUS AX024455 3924 bp DNA linear PAT 15-SEP-2000  
DEFINITION Sequence 2 from Patent DE19857750.  
ACCESSION AX024455  
VERSION AX024455.1 GI:10184623  
KEYWORDS human.  
SOURCE Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3924)

AUTHORS Stremmel, W.

JOURNAL Patent: DE 19857750-A 2 24-FEB-2000;

STREMMEL WOLFGANG (DE)

FEATURES Location/Qualifiers

1..3924

source /organism="Homo sapiens"

/db\_xref="taxon:9606"

/note="WDR3/menschliche MDR-Genfamilie"

BASE COUNT 1145 a 790 c 977 g 1012 t

ORIGIN

## Alignment Scores:

Pred. No.: 5,8e-146 Length: 3924  
Score: 1924.50 Matches: 372  
Percent Similarity: 74.10% Conservative: 123  
Best Local Similarity: 55.69% Mismatches: 162  
Query Match: 57.91% Indels: 11  
DB: 6 Gaps: 3

US-09-873-409-1 (1-659) x AX024455 (1-3924)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20  
Db 1851 GTAAATGTGGAGCAAGGAAGCACAGCAACTGATGAAGAGAGGGGTGACTTCAAA 1910  
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31  
Db 1911 CTTGTCAACATGCACACATCAGGAAGCCAGATCCAGTCAGAGAATTTGAACATAATGAT 1970  
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51  
Db 1971 GAAAGGTCGCCACTAGAAATGCCCAAAATGGCTGGAATCTCGCTATTTAGGCATTCT 2030  
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69  
Db 2031 ACTCAGAAAAACCTTAAAAAATTCACAAATGTGTCAAGAGAGCCTTGATGTGGAACCCGAT 2090  
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89  
Db 2091 GGACTTGAGCAAAATGTGCCACAGCTGCTTCTTGAAGGTCTCGAACTGAATAAACA 2150  
Qy 90 GluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109  
Db 2151 GAATGGCCCTACTTTGTCTGGGAACAGTATGTGCCATTTGCCAATGGGGGCTTCAGCG 2210  
Qy 110 ValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAspLysThrThr 129  
Db 2211 GCATTTTTCAGTCATATCTCAGAGATCATAGCAGATTTTGGACCAGCGCATGACAGTG 2270  
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149  
Db 2271 AAGCAGCAGAGTGCACATATTTCTTTGATTTTCTTATTTCTGGAAATATTTCTTT 2330  
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169  
Db 2331 TTTACTTTCTCTTCAGGGTTTCAGTTTGGAAAGCTGGCGAGATCTCTACCGAGA 2390  
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189  
Db 2391 CTGCGTCAATGGCTTTAAAGCAATGCTAAGACAGGACATGAGCTGTTGATGACCAT 2450  
Qy 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209  
Db 2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTGGCCACAGATGCTGCCCAAGTCCAAGA 2510  
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229  
Db 2511 GCCACAGGACAGGTGGCTTTAATGTGACAGAAATATAGCTTAACCTTGGAACTGGTATT 2570  
Qy 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249  
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAAACCTATTGCTATTAGCAGTTGTTCCAAAT 2630  
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269  
Db 2631 ATTGCTGTGTGAGGAATTTGTAATTTGAAATTTGGCTGGGAATTTGCCAAAAGAGATAAA 2690  
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289  
Db 2691 AAAGAACTGGAACTGCTGGAAGATTGCAACAGAGCAATAGAAAATATTAGGACAGTT 2750  
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309  
Db 3831 TTTTCAATGGTGTGTCAGGCT 3854

Db 2751 GTGTCTTTGACCCAGGAAAGAAATTTGAATCAATGATGTTGAAAAATTTGATGGACCT 2810  
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329  
Db 2811 TACAGGAATTTCTGTGCAGAGGCAACATCTATGGAATTTACTTTTAGTATCTCACAGCA 2870  
Qy 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349  
Db 2871 TTTATGATATTTTCTATGCGGTTGTTTTCGATTTGGTGCATATCTCATTTGTGAATGGA 2930  
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyValMetAla 369  
Db 2931 CATATGCCCTTCAGAGATGTTATTTCTGTGTTTCTGCAATTTGATTTGGTGCAGTGCT 2990  
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389  
Db 2991 CTAGGACATGCCAGTTTCATTTGCTCCAGACTATGCTAAAGCTTAAGCTGCTCTGCAGCCCAC 3050  
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409  
Db 3051 TTATTCATGCTGTTTGAAGAACAACTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110  
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429  
Db 3111 CCTGATAAATTTGAAGGAATATTAACATTTAATGAAGTCGTGTTCAACTATCCACCCGA 3170  
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449  
Db 3171 GCAAACTGTCAGTGTCTCAGGGCTGAGCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230  
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469  
Db 3231 CTGTGGGACAGTGGCTGTGGGAAGACGACGCGTGTGTCAGCTCTCGAGCGGTCTTAC 3290  
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489  
Db 3291 GACCCCTGGGGGACAGTGTCTCTGAGTCAAGAAAGCAAGAAACTCAATGTCCAG 3350  
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509  
Db 3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGAGCTATCTTATTGACTGCGAGCAT 3410  
Qy 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529  
Db 3411 GCCAGAAATATTCCTATGAGACAAACAGCGGGTGTATCACAGGATCAAAATTTGTGAGT 3470  
Qy 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysThrAsnThr 549  
Db 3471 GCAGCAAGCTGCCAACATACATCTCTTCATCAGAGCTTACCCCAAAATATGAAACA 3530  
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569  
Db 3531 AGAGTGGAGATAGGGAGCTCAGCTCTCAGAGGTCAAAAACAGAGGATTTGCTATTGCC 3590  
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAsp 589  
Db 3591 CGAGCCCTCATCAGACAACTCAAAATCCTCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650  
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609  
Db 3651 ACTGAAGTGAAGGTTGTCCAGAACCCCTGGCAAAAGCCAGAGAAGGGCGCCACTGCG 3710  
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629  
Db 3711 ATTGTGATTGCTACCGCTGTCCACCATCAGATGCGAGCTTAATAGTGTGTTTTCAG 3770  
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyr 649  
Db 3771 AATGGACAGTCAAGGACATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830  
Qy 650 PheLysLeuValAsnAlaGlnSer 657  
Db 3831 TTTTCAATGGTGTGTCAGGCT 3854



## RESULT 8

AX409652  
LOCUS AX409652 3924 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 2299 from Patent WO0229103.  
ACCESSION AX409652  
VERSION AX409652.1 GI:21442357  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1.  
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.  
Gene expression profiles in liver cancer  
Patent: WO 0229103-A 2299 11-APR-2002;  
GENE LOGIC INC (US)

## FEATURES

source Location/Qualifiers  
1..3924  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="EMBL/GenBank Accession No. M23234"

BASE COUNT 1145 a 790 c 977 g 1012 t

## ORIGIN

## Alignment Scores:

Pred. No.: 5.8e-146 Length: 3924  
Score: 1924.50 Matches: 372  
Percent Similarity: 74.10% Conservative: 123  
Best Local Similarity: 55.69% Mismatches: 162  
Query Match: 57.91% Indels: 11  
DB: 6 Gaps: 3

US-09-873-409-1 (1-659) x AX409652 (1-3924)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20  
DB 1851 GTAATTGTGGAGCAAGGAGCCACAGCACTGATGAAGAAGGAGGGGTACTTTCAA 1910  
QY 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31  
DB 1911 CTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAGAATTTGAATTAATGAT 1970  
QY 32 GluGlnMetGluSerMetTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51  
DB 1971 GAAAAGCTGCCACTAGAAATGGCCCAATGGCTGGAATCTCGCCTATTAGGCATTCT 2030  
QY 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69  
DB 2031 ACTCAGAAAACCTTAAATAATCACAAATGTGTCAAGAGAGCCCTTGATGGAAACCCGAT 2090  
QY 70 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 89  
DB 2091 GGACTTGAACCAATGTGCCACCAAGTCTCTTCTGAGAGTCTGAACTGAATAAACA 2150  
QY 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109  
DB 2151 GAATGGCCCTACTTTGTGGGGAACAGATATGTCATTCCTCAATGGGGGCTTCAGCCG 2210  
QY 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129  
DB 2211 GCATTTTCATGCATATTTCTCAGAGATCATAGCGATTTTGGACCGCGCATGATGAGTG 2270  
QY 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149  
DB 2271 AACGACAGAGTGCACATATTTCTTTGATTTTCTTATTTCTGGGAATATTTCTTTT 2330  
QY 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169  
DB 2331 TTTTACTTTCTCTTCAGGGTTTCAGCTTTGGGAAAGCTGGCGGAGATCTCCACCAAGA 2390  
QY 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189  
DB 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189

DB 2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATAGCTGGTTGATGACCAT 2450  
QY 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209  
DB 2451 AAAACAGTACTGTGTGCACTTTCTACAGACTTGCACAGATGCTGCCCAAGTCCCAAGA 2510  
QY 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229  
DB 2511 GCCACAGAAACCCAGGTGGCTTTAATTTGCACAGAAATATAGCTTAACCTTGGAACTGTATT 2570  
QY 230 IleIleSerPheIleTyrGlyTyrGluMetThrPheIleLeuLeuSerIleAlaProVal 249  
DB 2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCTTATGCTATTAGCAGTTGTTCCAAAT 2630  
QY 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269  
DB 2631 ATTGCTGTGCAGGAATGTTGAAATGAAATGTTGGCTGGAAATGCCAAAGAGATAAA 2690  
QY 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289  
DB 2691 AAAGAACTGGAAAGCTGCTGGAAGATTCACACAGAGCAATAGAAAATATTAGGACAGTT 2750  
QY 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309  
DB 2751 GTGCTTTTGACCCAGAAAGAAATTTGAATCAATGATATGTTGAAATTTGTATGACCT 2810  
QY 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329  
DB 2811 TACAGGAATCTGTGCAGAGGCACACATCTATGGAATTTACTTTTAGTATCTCACAGCA 2870  
QY 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349  
DB 2871 TTTATGATATTTTCTATGCCCGTTGTTTTCGATTTGGTGCATATCTCATTTGTAATGA 2930  
QY 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369  
DB 2931 CATATCGCGCTTCAGAGATGTTATTTCTGGTGTCTTCGAAATTTGATTTGGTCAGTGGCT 2990  
QY 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHis 389  
DB 2991 CTAGGACATGCCAGTTTCAATTTGCTCCAGCATATGCTAAAGCTTAAGCTGTCTCAGCCAC 3050  
QY 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409  
DB 3051 TTATTCATGCTCTTTTGAAGAACAACCTCTGATTGACAGTACAGTCAAGAGGGGCTGAAG 3110  
QY 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429  
DB 3111 CCTGATAAATTTGAAGGAATAATAACATTTAATGAAGTCTGCTTCAACTATCCACCCCA 3170  
QY 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449  
DB 3171 GCAACGCTGCCAGTGCTTCAGGGCTGAGCTGGAGGTGAAGAGGCCAGACATAGCC 3230  
QY 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469  
DB 3231 CTGTGGGCGAGCAGTGGCTGTGGGAAGACAGCGTGGTCCAGCTCTCTGGAGCGTCTTAC 3290  
QY 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489  
DB 3291 GACCCCTTGGCGGGGACAGTGGCTTCTCATGTTGTCAGAGCAAGCAAGAACTAATGTCCAG 3350  
QY 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509  
DB 3351 TGGCTCAGAGCTCACTCGGAATCGTGTCTCAGAGAGCTATCTCTATTTCAGTCGACATT 3410  
QY 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529  
DB 3411 GCCGAGAATATTTCCTATGGAGACAACACCGCGGTGTTATCATCAGGATCAAAATTTGAGT 3470  
QY 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549  
DB 3471 GCAGCCAAAGCTGCCAACATATATCTTTTCATCGAGAGCTTACCCCAACAATATGAACA 3530



```
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaLelela 569
:::|||||
Db 3531 AGATGGGAGATGAAGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTCCTATTGCC 3590
:::|||||
Qy 570 ArgAlaLeuLeuGlnLysProLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 589
:::|||||
Db 3591 CGAGGCCCTCATCAGCAACCTCAATCCCTCTGTTGGATGGAAGCTACATCAGCTCTGGAT 3650
:::|||||
Qy 590 AsnArgSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
:::|||||
Db 3651 ACTGAAAGTGAAGAGTTGTCAGAGNAGCCCTGGACAAAGCCAGAGAGCGCCGACCTGC 3710
:::|||||
Qy 610 LeuValValThrHisArgLeuSerAlaLeuGlnAsnAlaAspLeuLeuValValLeuHis 629
:::|||||
Db 3711 ATTGTGATTGCTCAGCGCTGTCACCATCCAGATCCAGACTTAAATAGTGGTGTTCAG 3770
:::|||||
Qy 630 ArgGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 649
:::|||||
Db 3771 AATGGGAGAGTCAAGGAGCATGGCAGCATCCAGCAGCTGCTGGCAGAGAAAGGCATCTAT 3830
:::|||||
Qy 650 PheLysLeuValAsnAlaGlnSer 657
:::|||||
Db 3831 TTTTCATGCTCAGTCTCAGGCT 3854
:::|||||

RESULT 9
HUMMDR3
LOCUS Human membrane glycoprotein P (mdr3) mRNA linear PRI 11-JUN-1993
DEFINITION M23234
ACCESSION M23234
VERSION M23234.1 GI:187501
KEYWORDS P-glycoprotein; membrane glycoprotein.
SOURCE Human liver, cDNA to mRNA, clone 3.27.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS van der Blik,A.M., Koolman,P.M., Schneider,C. and Borst,P.
TITLE Sequence of mdr3 cDNA encoding a human P-glycoprotein
JOURNAL Gene 71 (2), 401-411 (1988)
MEDLINE 89138016
PUBMED 2906314
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by P. Borst, 21-MAR-1989.
FEATURES
Source
1..3924
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..3924
/product="mdr3 mRNA"
33..3872
/note="P-glycoprotein"
/codon_start=1
/protein_id="AA36207.1"
/db_xref="GI:307181"
/translation="MDLEAAKNGTAMRPTSAEDGPELGTSQKREKTKTKVMIGVLT
LFRYSDWDLKPLMSLGTITMAIAGSGLLPLMMIVFEGMTDKFVDTAGNFSFPVNFSLSL
LNPGLLEEMTRYAVYVYSGLAGVLVAAYIQVSPMTLAAGROIRKIQKPFHAIHQ
EIGWFDINDTLENLRLDDISKISEGIDGKGVMFQAVAFAGFVGFIRGWKLTIL
VIMATSPILGUSAAMWAKILSAFSDKELAAAYAKAGAAVEALGAIRTVIAFGQNKEL
ERYQHLNKAKEIGIKKAIISANISMGIAFLIYASALAFWYGLTVISKEYTIGNAM
TVFFSILITGAFSGQAAPCIDAFANARGAAYVIFDIIDNNPKIDSFSERGHKPSIKG
NLEFNDVHFSYPSRANVKILGLNLKQSGVTVALVSGSGCKSTTVOLIORLYDPDE
GTINIDGQDIRNFNVNLYREILIGVVSQBPVLEFSTTIAENICYGRGNTMDERKKAVKE
ANAYEFIMKLPKQFTLVERGAQUSGQKQRIARALVRNPNKILLDLDEATISALDTE
SEAEVQALDDKAREGRTTIVIAHRLSTVRNADVIAGFEDGVIVEGSHSELMKKEGVY
FKLVNMQTSQSGIQSEEFELNDEKATMAPNGMKSRLFSQTNLKNQSMQCKSLD
VETDLEARNVPVSLKLMNKTWPYFVVGTCVCAIANGGLQAPVIFSEIIAIFG
PGDDAVKQKQKNCI FSLIFLGI ISFTFFLQGTFGKAGEILLTRLRSMAPKMLRQ
DMSVDFDHTNGSLSTRLADAAQVQAGTQFRLALIAQNTIANLGTIIISFIYQWL
TULLLAVPIITAVSGIVEMKLIAGNAKDKKELEAGKIATIAENIRNTVSLTQERK
FESMTVEKLYGYSVQKAHITYGTFSIQAPMYFSYAGCFRFGAYLIVINGHMFRD
VILVPSAIVFGAVALGHASSFAPYAKAKLSAAHLFMLFERQPLIDSYSEELKPKDF
```

```
EGNITFNEVVPYPTANVPVLOGLSLEVKGQTLALVSGSGCKSTVQLBERFPDP
LAGTVLDGGEAKKLINQWLRAGQIGIVSQPILFDCSIAENIATIGDMSRVVSQDEIYS
AAAXLHDPHPIETLPHKYETRGDKQRIAIARALIRPOILLDDATSA
LDTSESVQOALDKAREGRTICVIAHRLSTIQNLADLIVFQNGRVEHGTQOULLAQ
KGIYFSMVSVOAQTNL"
BASE COUNT 1145 a 790 c 977 g 1012 t
ORIGIN
Alignment Scores:
Pred. No.: 5,8e-146 Length: 3924
Score: 1924.50 Matches: 372
Percent Similarity: 74.10% Conservative: 123
Best Local Similarity: 55.69% Mismatches: 162
Query Match: 57.91% Indels: 11
DB: 9 Gaps: 3
US-09-873-409-1 (1-659) x HUMMDR3 (1-3924)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
:::|||||
Db 1851 GTAAATTGTGGAGCAAGGAAGCCACAGCAACTGATGAAGAAGGAAGGGGTGACTTCAAA 1910
:::|||||
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31
:::|||||
Db 1911 CTGTCAACATGCACACATCAGGAAGCCAGATCCAGTCAGAGAATTTGNACTAATGAT 1970
:::|||||
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
:::|||||
Db 1971 GAAAGGCTGCCACTAGAAATGCCCAAAATCGCTGAAATCTCGCTATTTAGGCATTCT 2030
:::|||||
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
:::|||||
Db 2031 ACTCAGAAAAACCTTAAAAAATTCACAAATGTGTCAAGAAGAGCCTTGATGTGGAAACCGAT 2090
:::|||||
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
:::|||||
Db 2091 GGACTTGAAGCAAAATGTGCCACACAGTGTCTTCTTGAAGGTCTGAAACTGAATAAACA 2150
:::|||||
Qy 90 GluTrpPropheValIleuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
:::|||||
Db 2151 GAATGGCCCTACTTGTGCTGGGAACAGTATGTGCCAATGGGGGGCTTCAGCGC 2210
:::|||||
Qy 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
:::|||||
Db 2211 GCATTTTTCAGTCATATCTCTCAGAGATCATAGCGATTTTGGACACAGGCGATGATG 2270
:::|||||
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
:::|||||
Db 2271 AAGCAGCAGAGTGTCAACATATCTCTTGTATTTCTTATTTCTGGAAATATTTCTTTT 2330
:::|||||
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
:::|||||
Db 2331 TTTACTTCTCTTCAGGGTTTCAGTTTGGGAAAGCTGGCGAGATCTCCACAGAAGA 2390
:::|||||
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
:::|||||
Db 2391 CTGCGCTCAATGGCTTTTAAAGCAATGTCTAAGACAGGACATGAGCTGTTGATGACCAT 2450
:::|||||
Qy 190 GluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
:::|||||
Db 2451 AAAAAACAGTACTGGTGCACTTCTTCAAGACTTGCACAGATGTGCCCAAGTCCCAAGA 2510
:::|||||
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
:::|||||
Db 2511 GCACAGGACACAGGTGGCTTTAATTCACAGATATATAGCTTAACCTTGGAACTGGTAT 2570
:::|||||
Qy 230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerIleAlaProVal 249
:::|||||
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCCTATTGCTATTAGCAGTTGTTCCAAT 2630
:::|||||
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
:::|||||
Db 2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTTGTTGGCTGGAAATGCCCAAGAGATAAA 2690
:::|||||
```





ORGANISM *Cricetulus griseus*  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Cricetulus.

REFERENCE 1 (bases 1 to 4296)  
Endicott, J. A., Sarangi, P. and Ling, V.  
Complete cDNA sequences encoding the Chinese hamster P-glycoprotein  
gene family  
JOURNAL DNA Seq. 2 (2), 89-101 (1991)  
MEDLINE 92135896  
PUBMED 1685679

FEATURES  
Location/Qualifiers  
1..4296  
/organism="Cricetulus griseus"  
/db\_xref="taxon:10029"  
/tissue\_type="liver and ovary"  
/dev\_stage="adult"  
103..3933  
/note="putative"  
/codon\_start=1  
/product="p-glycoprotein isoform I"  
/protein\_id="AA68883.1"  
/db\_xref="GI:191165"  
/translacion="MEFEEDPSGRKDNFLKMGKRSKKEKKEKPPVSVVFTWFRYAGW  
LDRLMYLGLTAAIIGHVADPLMVLFGDMTDSFASVGNIPFNATNNAQVNASDIFG  
KLEEMTYAYTYTGAGQVLIIVAYIOVSFCLAAGQIHKIRQKFFHAIIMNOIGWF  
DVHDVGLNRLTDDVSKINGIDGKIMFQMAWTFGGPIIGFTGKWLTLVILAI  
SPVLGASGIWAKILSFTDKLOAYAKAGAVEVLAAITVIAFGOKKELRYNN  
NLEBAKGLIKAITANISMGAAPELLIYASVALAFWGTSLVLSKEVSGOVLTVFPA  
VLIGAFSIGQAPNIEAPANGAAYEIFNIIDNKPSLDSFKNGYKPDNIGNLEPK  
NIHFSYRDKQVILKGLNVKQSVQVVALVNGSCGKSTTVQLLRLYDPTGVVSI  
DGODIRTNVRLREIIGVWSQEPVLPAITAEINRYGRENVTWDEIEKAVKEANAYD  
FIMKLPHKPTDLVGERGAOLSGGOKORIAIARALVRNPKILLLDEATSSALDTESEAVV  
QAALDKAREGTTVIAHRLSTVRNADLIAGFDGVLVEQGNHLEKREGLIYKLVN  
TQAGNIELGNEVESKNEIDNLDMSKDSASSLIRRRSIRHNGDQDKLSTK  
EALDEDPPIISFWRLKLNSENSEFYVVGIFCAIVNGALQPAFSIIISFKVGVFTRNT  
DDETRHDSNLSILLFIIIVISPTFFLQGTFFKAGEILTKELRYVFKSMRLQDV  
SWFONKNTGALTTRLANDAGQVKATGARLAVITONIANLGTGIIISLIYQWLT  
LLLAIVPIATAGVVEKMWLSQALKDKKELEGSGKIATEAENFRVTVSLTRRQKPE  
NMYAQSLQIPYRNALKKXHVGTIFSTQAMMYESYACFPFGAYLVARELMTENVL  
LVFSALVFNAMVQSSFPADYAKAVKASASHIIMIEKPSIDISYSTGGKPNLTLEG  
NVKFNVEVNTPTDPIVGLNLEKVGKQTLALVSSGCGKSTVQLLEFVDPDMA  
GTVPDLQFVNLQWMLRHGIVSQEPILFDGSIATENAYGNRSVVSQDETERAA  
KEANIDHGFIEQLPKYTRVGDQTLSSGGOKORIAIARALVRPHILLLDEATSSALD  
TESKVVQEALDKAREGRTCTVIAHRLSTIQNADLIIVIQNGKVEHGHQQLLAQKG  
IYFSNVSVQAGAKR"

BASE COUNT 1271 a 892 c 1029 g 1104 t

Alignment Scores:  
Pred. No.: 3,19e-145 Length: 4296  
Score: 1916.00 Matches: 360  
Percent Similarity: 74.96% Conservative: 143  
Best Local Similarity: 53.65% Mismatch: 154  
Query Match: 57.66% Indels: 14  
DB: 10 Gaps: 4

US-09-873-409-1 (1-659) x CRUGPI (1-4296)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
|||||  
|||||  
Db 1906 GTCATTGGGAGCAAGAAATCATGAGAGCTCATGAGAGAGAGGCGATTACTTCAA 1965  
|||||  
Qy 21 LeuValMetSerGln-----AspIleLysLysAla 30  
|||||  
Db 1966 CTGTGTCATGACACAGACAGCAAGAAATGAAATGAAATGAGAAATGAGTGTGAGTCT 2025  
|||||  
Qy 31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48  
|||||  
Db 2026 AAAAAATGAAATGTAATTTAGCATGCTCTCAAAGAGATTACAGCATCCAGTCTAATCAGA 2085  
|||||  
Qy 49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68  
|||||  
Db 3166 GCCTCTGAAGCCCTAAATACATTTGAAGAAATGTGAATTTAAATGAAGTCGTGTTTCAACTAC 3225

UUS-09-873-409-1 (1-659) x CRUPEGPI (1-4296)

Qy	1	MetLeuAlaGluLysGlyAlaHisIaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
		::::: ::::  ::    ::    ::    ::    ::    ::    ::    ::    ::	
Db	1906	GTCATTGTGGCAAGGAATCATGAAGAGCTCATGAGAGAAGGCATTTACTTCMAA	1965
		::::: ::::  ::    ::    ::    ::    ::    ::    ::    ::    ::	
Qy	21	LeuValMetSerGln-----AspIleLysLysAla	30
		::    ::    ::    ::    ::    ::    ::    ::    ::    ::	
Db	1966	CTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGCAAATGAAGTTGGTGAGTCT	2025
		::::: ::::  ::    ::    ::    ::    ::    ::    ::    ::    ::	
Qy	31	AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----	48
		::::: ::::  ::    ::    ::    ::    ::    ::    ::    ::    ::	
Db	2026	AAAAATGAATTCATAATTTAGACATGCTCTTCAAAGAGATCCAGCATCCAGCTTAATCAGA	2085
		::::: ::::  ::    ::    ::    ::    ::    ::    ::    ::    ::	
Qy	49	ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr	68
		::: ::::  ::    ::    ::    ::    ::    ::    ::    ::    ::	

Qy	427	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleClnuArgGlyLys	446
Db	3256	CCACCCGGCTGACATCCAGTGTCTCAGGGCTGAACCTGGAGGTGAAGAGGCCAG	3285
Qy	447	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln	466
Db	3286	ACGCTGGCCCTGGTGGGACGAGTGGCTCGCGGAAGAGACACAGTGTCCAGCTGCTCCAG	3345
Qy	467	ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu	486
Db	3346	CGGTCTACGACCCATGGCCGGAACAGTGTCTAGATGGCAAGAGTAACCACTA	3405
Qy	487	AsnValGlnTrpLeuArgSerGlnIleAlaValProGlnClnuProValLeuPheAsn	506
Db	3406	AATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCAGAGCCCATCTGTTTAC	3465
Qy	507	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	526
Db	3466	TGCAGCATCGCTGAGAACATCGCTTATGGAGAACACGCGGGTCTGTGCCAGGATGAG	3525
Qy	527	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
Db	3526	ATCGAGAGGGCGGCGAAGAACCAATCCACAGTTCATTGAGTCACTGCTGCACAAA	3585
Qy	547	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
Db	3586	TACACACACAGAGTAGGACACAAAGGACTCAGCTGTCGGTGGGCGAGAACACGCATC	3645
Qy	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspClnuAlaThrSer	586
Db	3646	GCCATAGCACGCGCCCTGTCAGACAGCTCACAATTTGCTCCTCGATGAAGCAACATCA	3705
Qy	587	AlaLeuAspAsnAspSerClnuLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
Db	3706	GGCCTGGATACAGAGAGTGAAAGGTGTCCAGGAGCTCTCGACAAAGCCGAGAGGCG	3765
Qy	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
Db	3766	CGCAGCTGATGTGATCGCTCACCGCTGTCCACCATCCAGACGCGAGCTTGATCGTG	3825
Qy	627	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg	646
Db	3826	GTGATTCAGATGGCAAGTCAAGGAGTGCACCCACCCAGCAGCTGTGTCACAGAAA	3885
Qy	647	AspIleTyrPheLysLeuValAsnAlaGlnSer	657
Db	3886	GGCATCTATTCTCCATGTCAGTGTGCAGGCT	3918
RESULT	12		
LOCUS	CRUGP1165	4304 bp mRNA linear ROD 27-APR-1993	
DEFINITION	Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.		
ACCESSION	M59253		
VERSION	M59253.1	GI:191154	
KEYWORDS	multidrug resistance glycoprotein; p-glycoprotein.		
SOURCE	Chinese hamster, CDNA to mRNA, clone ADX165.		
ORGANISM	Cricetulus sp.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;		
	Cricetulus.		
REFERENCE	1. (bases 1 to 4304)		
AUTHORS	Devine, S.E., Huesain, A., Davide, J.P. and Melera, P.W.		
TITLE	Full length and alternatively spliced pgpl transcripts in		
	multidrug-resistant Chinese hamster lung cells		
JOURNAL	J. Biol. Chem. 266 (7), 4545-4555 (1991)		
MEDLINE	91154265		
PUBMED	1671863		
FEATURES	Location/Qualifiers		
source	1..4304		
	/organism="Cricetulus sp."		
	/db_xref="taxon:10031"		
	/cell_type="fibroblast"		

		/tissue_type="lung"	
		1..4304	
		/gene="pgp-1"	
		111..3941	
		/genes="pgp-1"	
		/codon_start=1	
		/product="p-glycoprotein"	
		/protein_id="AAA37004.1"	
		/db_xref="GI:191155"	
		/translation="MEFEEDFSGRKDKNFKMGKSKKEKKEKPKVVSVFTMFRYAGW	
		LDRLMLVGTAAALIHGVALPLMLVFGDMTDFASVGNITPTNATNATONVNASDIFG	
		KLEEMTYAYVYGTIGAGVLIVAYIOVSFWCLAAAGROJHKIROKFEHAINNQELGW	
		DVHDVGLNRLTDDVSKINEGIDGKIMFEQMATFFGFIIGFTGKWLTLVILAI	
		SPVLGSAGIWAAILSSFTDKELQAYAKAGAAVEELAAIRTVIAFGGQKELSRINN	
		NLEAEKRLGIKKAITANISGAAPFLIYAYALAFWYGTSLIVSKYSIGQVLTFFFA	
		VLTAPSIGQASNIEAFANARGAAEYFNIIDNKPSIDFSKNGKYPDNIKGLEFK	
		NIHFSYPRKDVOILKGLNLKVGSGQVVALVNGSGCGKSTTVQLQLRYDTEGVVSI	
		DGQDITINRYIREIIGVVSORPVLPAITIAENIRYGRNVTWDEIEKAKEANAYD	
		FIMKLPHKPTLHGERGAQLSGGOKORIAARALVRNPKILLDLDEATSALDTESEAVV	
		QAALDKAREGRTTIVIAHRLSTVRNADIAGFDGGVIVEQGNHEELMREKGIYFKLVV	
		TQAGNEIELNGEVESKNEIDNMDSSKDSASLIRRRTRRSIRPHDPDRKLSKT	
		EALDEDVPPTISFWRLILKINSWPYFVVGIFCAIVNGALQAPAFSIIIFSKVGVFTRNT	
		DDETKRHSNLSLPLILGVISFTTFQGFPGKAGEILTKRLRYMFKSMRLQDV	
		SWFDPNKTTCALITRLANDAGQVKGATGARLAVITONIANLGTGIIISLYGQHLT	
		LLLAIVPIIAGVVMKLSGQALKDKKELEGSKTATEIAENFRVTSLTRQKFE	
		NMTAQSLQIPYRNALKKARVFGITFSFTQAMTFYSAAACFRFGALVARELMTFENVL	
		LVFSASIVFGAMVGQSVSSFADYAKAKVSASHIIMIIEKVPISIDYSTGGLKPNLLEG	
		NVAFNEVFNYPTRPDIPLVGLNLVKGKQTLALVSGSGCGKSTVQLLERFYDPMA	
		GTPLDGKEYNOLNOLRAHLGIVSOEPIILFDCSIAENIAYGNSVSVQDEITERAA	
		KEANIHOFIESLDPKNTRVGDGKQSLSGGOKORIAARALVRPHILLDLDEATSALD	
		TESEKVVQEAALDKAREGRTCTVIAHRLSTIQNALDIVIYQNGKVEHGTGQLLAQKG	
		IYFSMVSVQAGAKR"	
BASE COUNT	1275 a	895 c	1031 g 1103 t
ORIGIN	Map position 1q26.		
Alignment Scores:			
Pred. No.:	3,19e-145	Length:	4304
Score:	1916.00	Matches:	360
Percent Similarity:	74.96%	Conservative:	143
Best Local Similarity:	53.65%	Mismatches:	154
Query Match:	57.66%	Indels:	14
DB:	10	Gaps:	4
US-09-873-409-1 (1-659) x CRUGP1165 (1-4304)			
Qy	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer	20
Db	1914	GTCAATGTGGAGCAGGAAATCATGAAGAGCTCATGAGAGAGGCGCATTTACTCAA	1973
Qy	21	LeuValMetSerGln-----AspIleLysLysAla	30
Db	1974	CTTGTGATGACACACAGCAGGAAATGAAATTGAATTAGGAATGAAGTTGGTGA	2033
Qy	31	AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----	48
Db	2034	AAAAATGAAATTTGATAATTATGACATGTCTTCAAAAGATTCAGCATCCAGTAA	2093
Qy	49	ProLeuHisSerValLysSerLysSerAspPheIleAspLysAlaGluGluSerThr	68
Db	2094	AGGAGATCAACTCGCAGAAAGTATCCGTGGACCATGACCAAGCAGGAAGCTTAG	2153
Qy	69	---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	87
Db	2154	AAAGAGGCTTGGATGAAGATGACTCCATTTCTTTGGCGTATCTCTAAAGTTGAT	2213
Qy	88	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	107
Db	2214	TCATCTGAATGGCCATATTTTGTGTTGTGTATATTTCTGTGCGCATAGTAAAT	2273
Qy	108	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsn	126
Db	2274	CAACACGATCTTCAATAATATTTTCTAAGGTGTGAGGGTTTTTCAAGAAATACTGAT	2333

QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146  
 DB 2334 GATGAACCAACAGCATGATAGCACTTGTGTTCTTCTTATTTCTGATCCTTTGGGTC 2393  
 QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166  
 DB 2394 ATTCTTTTATACATTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGATCCTC 2453  
 QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186  
 DB 2454 ACCAAGCGACTCCGATACATGTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTT 2513  
 QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGln 206  
 DB 2514 GATAACCTTAAACACCACTGGAGCATTTGACCACAGGCTCGCCCAACGATGCTGTCAA 2573  
 QY 207 IleGlnGlyValThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
 DB 2574 GTTAAAGGGCTACAGAGCCAGACTTGTGTCATTTACCAGATATAGCAATCTTGGG 2633  
 QY 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246  
 DB 2634 ACAGGAATCATCATATCCCTTAATCTATGCTGGCAGTTGACACTTCTACTCTTAGCAAT 2693  
 QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
 DB 2694 GTCCCCATCATTCAGCAGGAGTGGTTGAGATGAAATGTTGTCTGCAACGACTA 2753  
 QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
 DB 2754 AAGATTAAGAGGAGCTAGAGGTTCTGGAGATTTGCTGTAAGCAATAGAGAACTTC 2813  
 QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
 DB 2814 CGCACTGCTCTCTTTGACTCGGAGCAGAGCTTTGAAATATGATGCCAGACCTTG 2873  
 QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
 DB 2874 CAGATACCATACAGAAATGCTCTGAAGAAAGCAGACAGCTCTTTGGGATCATCTCTCTTC 2933  
 QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346  
 DB 2934 ACCAGGCCATGATGATTTTTCGTATGCTGCTGTTTTCGGTTGGCTTACTTGGTG 2993  
 QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
 DB 2994 GCACGGGAATTAATGACATTTGAAATGTTCTATTAGTATCTCAGCTATGTTCTTGT 3053  
 QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
 DB 3054 GCATGGCAGTTGGGAGGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAAGTGTG 3113  
 QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
 DB 3114 GCATCCCATATCATGATGATCATTTGAAAGTCCCTTCCATTGACAGCTACGACCGGA 3173  
 QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuLeuPheArgGluValSerPheTyr 426  
 DB 3174 GGCTGAAGCCTTAATACATTTGGAAGGAATGTGAATTTAATGAAGTCGTTCACATAC 3233  
 QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
 DB 3234 CCCACCCGGCTGACATCCAGCTTCAGGGCTGAACTCGAGGTGAAGAGGGCCGAC 3293  
 QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
 DB 3294 ACCTGGCCCTGTGGGACGAGTGTGCTGGGAAGAGCAGCAGTGTGCTGCTGCTGAG 3353  
 QY 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
 DB 3354 CGGTTCTACAGCCCATGCCGGAACAGTGTCTTAGATGGCAAGAAAGTAAACCACTA 3413  
 QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506

DB 3414 AATGTCAGTGGCTCCGAGCACACCTAGGCATGTGTCCCAAGAGCCCATCTGTTTGAC 3473  
 QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526  
 DB 3474 TCGAGCATCGCTGAGACATCGCTTATGGAGACAACAGCCGGTGTGTGCCAGGATGAG 3533  
 QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
 DB 3534 ATCGAGAGGGCGCGCAAGGAAGCAACATCCACAGTTTCATTGAGTCACTGGCTGCACAA 3593  
 QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
 DB 3594 TACAACACCCAGAGTAGGAGACAAGAGGACTCAGCTGTGGGTGGGAGAAACAACGCATC 3653  
 QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586  
 DB 3654 GCCATGACAGCGCCCTCGTCAGACAGCTCACTTTGCTCTGGATGAAGCAACATCA 3713  
 QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
 DB 3714 GCCCTGGATACAGAGAGTGAAGGTTGTCCAGGAAGCTCTGGCAAGCCAGAGAGGC 3773  
 QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
 DB 3774 CGCACTGCATTTGATCGCTCACCGCTGTCCACCATCCAGAACCCAGACTGTGTCGTG 3833  
 QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646  
 DB 3834 GTGATTCAGATGCGCAGGTCAAGAGGATGGCACCCACCCAGCAGCTGTGTCGCACAGAA 3893  
 QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657  
 DB 3894 GGCATCTATTCTTCATGCTCAGTGTGCGAGCT 3926  
 RESULT 13  
 MUSMDRA  
 LOCUS  
 DEFINITION Mouse mdr gene encoding a multidrug resistance protein mRNA,  
 complete cds.  
 ACCESSION J03398  
 VERSION J03398.1 GI:199109  
 KEYWORDS multidrug resistance protein.  
 SOURCE Mouse (strain BALB/c) pre-B-cell (cell line 702/3), cDNA to mRNA,  
 clone MDR2.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 4084)  
 AUTHORS Gros, P., Raymond, M., Bell, J. and Housman, D.  
 TITLE Cloning and characterization of a second member of the mouse mdr  
 gene family  
 JOURNAL Mol. Cell. Biol. 8 (7), 2770-2778 (1988)  
 MEDLINE 88302195  
 PUBMED 3405218  
 REFERENCE 2 (sites)  
 AUTHORS Raymond, M. and Gros, P.  
 TITLE Mammalian multidrug-resistance gene: correlation of exon  
 organization with structural domains and duplication of an  
 ancestral gene  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (17), 6488-6492 (1989)  
 MEDLINE 89367274  
 PUBMED 2570420  
 REFERENCE 3 (sites)  
 AUTHORS Raymond, M. and Gros, P.  
 TITLE Cell-specific activity of cis-acting regulatory elements in the  
 promoter of the mouse multidrug resistance gene mdr1  
 JOURNAL Mol. Cell. Biol. 10 (11), 6036-6040 (1990)  
 MEDLINE 91042535  
 PUBMED 2248681  
 COMMENT Submitted in computer readable form by P.Gros 08-JUL-1988.  
 FEATURES Location/Qualifiers  
 source 1..4084



```

/organism="Mus musculus"
/strain="BALB/c"
/sub_species="domesticus"
/db_xref="taxon:10090"
/cell_line="702/3"
/cell_type="pre-B cell"
194..4024
CDS
/codon_start=1
/product="multidrug resistance protein"
/protein_id="AAA39516.1"
/db_xref="GI:387428"
/translation="MDLEARNGTARLDGDFELGSLISNOGRKKKKVNLIGLTLFR
YSDQKQFLMFLGTLMAIAGSGPLMMIVFEGMTDFDNTGNFSLVNFSLUSMLNP
GRILEEMETAYYYSLGGVLAAYIQVSWTAAQRQIKKIRKQKFFHAILQEMQ
WFDIKGTTELRLTDDVSKISBGIKGVGMFFQAIATFPAGFIVGFIKWKLTVM
AISPILGLTAVAKILSTSDKELAAKAGAVEAPCAIRTVIAFGQNKELRY
QKLELNKKTGIKKAISANISMGIAFLIJYASVALAPWGSTLVIKSEVTIGNMTVF
FSLIUGASVQAAPCIDAFANARGAAVIFDIIDNNPKIDSFSEGRKPDNTKGLNLE
FSDVHFSYPRANIKILKGLNKLKVKSGTVALVNGSCGKSTTVQLLRLYDPEGKI
SIDQDIRNFNRLREIIGVSOBEPVLFSTIAENIRYGRNVMTDEIEKAVEANA
YDFIMKLPQKFDLVGRGAQLSGQKQRIAIARALVRNPKILLDEATSDLTESA
EVOAALDKAREGRTTIVIAHRLSTIRNADVIAGFDGVIVEQGSHELMKKEGIYFRL
VNMOTAGSOILSEFEVELSDEKAAGDVA PNGWKARFRNSTKSKLSKSPHONLDEBT
NELDANVPVPSPLKVLKNTPEYFVGVTCALANGALOPAFSILLSEMIALFGPD
DAVKQCKNMFSLVLGLVLSFPTFLQGTFGKAGEILTLRLRSWAFKMLRQDMS
WPDHKNSTGALSTRLATDAQVQGTGKTLALIAQNTANLGTGIIISFYGQTLTL
LLSVPPVIAVAGIVEMKAGNKRDKMEAKIAEAIENIRTVSVISIQBRKTES
MYVEKLVGRNRSVRKHAHYIGITFSISOAFWFSYAGCFRFGSYLI VNGHMRKQVIL
VFSNVLGVALVNSKAPDYAKAKLSAAYLPSLPEROPLIDSYSGEGILWPKDFGS
VTNEVFNTPTRANVPVLOGLSLVKVKGQTLALVSGSGCGKSTVQVLLERFDPNAG
SVLLDQGEAKNQLWRAGIGVSOBEPVLFSTIAENIRYGRNVMTDEIEKAVEANA
EAMNHPPIETLPKYNTRVGDKGTLQSGQKQRIAIARALRQPRVLLLDENATSDLT
ESEKVOEALDKAREGRTTIVIAHRLSTIRNADVIAGFDGVIVEQGSHELMKKEGIYFRL
YFSMNVNQAGTQNL"
BASE COUNT 1117 a 934 c 1075 g 958 t
ORIGIN 43 bp upstream of HinfI site.
Alignment Scores:
Pred. No.: 3,27e-145 Length: 4084
Score: 1915,50 Matches: 366
Percent Similarity: 73,95% Conservatism: 128
Best Local Similarity: 54,79% Mismatches: 163
Query Match: 57,64% Indels: 11
DB: 10 Gaps: 1
US-09-873-409-1 (1-659) x MUSMDRA (1-4084)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 2003 GTCATTGTGGAACAAGGACGACAGTGTGATGAAAGAGGAGGATCTACTTCAGA 2062
Qy 21 LeuValMetSerGln-----AspIleLysLys 29
Db 2063 CTCGGTAAATGACACAGCAGAGGACGATCTGTGCAGAAAGATTGGAAGTTGAGCTA 2122
Qy 30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49
Db 2123 AGTGACGAAAGGCTGCTGGAGATGTGCCCAATGGCTGGAAGACGACGATATTAGG 2182
Qy 50 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluSerThrGln 69
Db 2183 AATTCTACAAAGAAAGCTTAAAGTCCACATCAGATAGGCTGAGTGAAGAAACCAAT 2242
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db 2243 GAACTTGATGCAACCGTCCACCGATGTCTTTCTGAAGGTCTTAAACTGAAATAACA 2302
Qy 90 GluTTPProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db 2303 GAGTGGCCCTACTTTGTGGTGGAAACAGCTCTGTGCATTCGCAATGGAGCCCTCCAGCG 2362
Qy 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129

```

```

Db 2363 GCTTCTCCATCATCTCTGTCTGAGATGATAGCTATCTTTGGCCCTGGGATGACGCAGTG 2422
Qy 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db 2423 AAGCAGCAAAAGTGTAACATGTTCTCCCTGCTCTTCTGGCCCTAGGAGTCTCTCTCTTC 2482
Qy 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db 2483 TTTACTTTCTCTCTCCAGGCTTCACGCTTTGGGAAAGCTGGAGAGATCCTCACCACAGG 2542
Qy 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTTPPheAspGluLys 189
Db 2543 CTCGGTCCATGGCTTTAAGCATGCTAAGCAGCAGCATGAGCTGGTTGATGATCAT 2602
Qy 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db 2603 AAAAACAGTACTGAGCAGCTTTCTACAAGATCTGCCACAGATGTCGCAAGTCCAAGGA 2662
Qy 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db 2663 GCCACGGGAACCAAGTTGGCTTTAATTGGCAGCAACACAGCAACCTTGGAAACCGGTATT 2722
Qy 230 IleIleSerPheIleTyrGlyTTPGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db 2723 ATTATATCATTTATTTAGGTTGGCACTGACACTTCTGCTGTTATCGTTGTTCCATTC 2782
Qy 250 LeuAlaValThrGlyMetIleGlyThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db 2783 ATTGCTGTAGCAGGAATTTGTTGAATGTTGGCTGGCAATGTCGCAAGAGAGATAAA 2842
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db 2843 AAGAAATGGAAGCTGCTGGGAAGATTGCAACAGAGCAATAGAAAAATATTTCGAACCTGT 2902
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db 2903 GTATCTTTGCCCAAGAAAGAAATTTGAGTCAATGATGTTGAAATAATTCATGCGACT 2962
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db 2963 TACAGGAATTCGGTGGCGAAGCAGACATCTACGGCATCACTTTTAGCATCTCCCAAGCA 3022
Qy 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db 3023 TTCATGTATTTTCTTATGCTGGCTGTTTTCGATTGCTTACCTAATTTGTGAATCGA 3082
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyValAlaMetAla 369
Db 3083 CATATGGCTTCAAGATGTCATTTCTGCTCTTTTCTGCATTTGCTTGGCGCGGTGCT 3142
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db 3143 CTAGACACGCCAGCTCATTTGCTCCGACATATGCAAAAGCCAAAGCTGCTGTCAGCATAC 3202
Qy 390 LeuPheAlaLeuLeuGlyLysPheAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db 3203 TTGTTTACGCTGTTTGAAGAACCACTCTGATTGACAGCTACAGTGGAGAGGCGGTGTGG 3262
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db 3263 CCTGATAAGTTTGAAGGAAGCGTGACATTTAATGAAGTCTGTTTCACTATCCACCCCGG 3322
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db 3323 GCCAACGTGCCAGTCTTCAGGGCTGAGCCTTGAGGTGAAGAGGGCCACAGCGTGGCC 3382
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db 3383 CTGTGGGCGAGCAGTGGCTGCGGGAAGAGCAGAGTGGTCCAGTCTGCTCGAGCGCTTCTAT 3442
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db 3443 GACCCCATGGCTGATCATGCTCTTAGATGTTCAAGAAAGCAAGAAATCAATTCATTCAG 3502

```





```
Qy 127 LysThrThrLeuLysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyVal 146
   :::::::::::::::::::::
Db 2365 CAGAAAGTTTTCGAGAAAGAACCTATATCTCACTGCTGTTTTTAGCACCTGGGATC 2424

Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
   :::::::::::::::::::::
Db 2425 ATTTCCTTTTATTACTCTCTTTGTTGAGGGTTTTGCAATTTGGCAAGCTGGAGAAATCTT 2484

Qy 167 ThrMetArgLeuArgHisIleAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePhe 186
   :::::::::::::::::::::
Db 2485 ACAATGAAGCTTCGATTCATGGCAATTAAGCAATGCTGAGACAGGACATGGCCTGGTTT 2544

Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
   :::::::::::::::::::::
Db 2545 GATGATCTCAAAAATAGCACTGGAGCACTGACAAACAAGACCTTGCTTAATGATGCTCACA 2604

Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
   :::::::::::::::::::::
Db 2605 GTAAAGGAGCAACTGGTGTGAGATTGGCATTAATTTGCCAAACATAGCTAAACCTGGGA 2664

Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
   :::::::::::::::::::::
Db 2665 ACTGGAATATTATATCTTTAGTTTATGCTGGCAGCTGACCTGTTACTCTTAGCTGTT 2724

Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
   :::::::::::::::::::::
Db 2725 GTACCAATCATTTGCAGTGGCAGGAATGATTGAATGAAGATGTTTGGCTGGGCACAAA 2784

Qy 267 LysAspLysGlnGluLeuLysHisIleGlyValIleAlaThrGluAlaLeuGluAsnIle 286
   :::::::::::::::::::::
Db 2785 AAAGATAAAATAGAGCTGGAAGCTGCAGGAAGATTTGCCACAGAGCCATAGAAAATATT 2844

Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
   :::::::::::::::::::::
Db 2845 CGAACTGTGCTAGTTTGACTTCGGGAAAAGATTTTGACTCATGTACGGAGAACATTG 2904

Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
   :::::::::::::::::::::
Db 2905 CTTGTACCATACAGAAATTTGTGTAAGAAGCGGCACATATTGGGTTTTGTTTTCCCTT 2964

Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
   :::::::::::::::::::::
Db 2965 TCACAAGCAATGATGTTTTTACCTATCTGCTGCTGTTTTCCGCTTTGGAGCCCTATCTTG 3024

Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
   :::::::::::::::::::::
Db 3025 GTTAATGACACATAGACTAGACAAAATCTGTTTTTGTAGTTTTCAGCTGTGTATTGTT 3084

Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
   :::::::::::::::::::::
Db 3085 GCTATGGCATTAGGACAAACACAGCTCTTTTGCACAGACTATGCCAAAGCCAAAGATATCA 3144

Qy 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
   :::::::::::::::::::::
Db 3145 GCAGCCCACTTGTGTTGCTGTTTAATAGAGTACCTCCAAATAGACAGATTACCGCAGGAT 3204

Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
   :::::::::::::::::::::
Db 3205 GGAGAGAACCTGAAATAATTTGGGGGAAACACAAGATCAAAAGATGTGAATTTAACTAT 3264

Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
   :::::::::::::::::::::
Db 3265 CCAACCCAGCAGAGGTCAAAATCTCCAGGCTTGAATCTAGCAGTAGAGAAAAGAGNA 3324

Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
   :::::::::::::::::::::
Db 3325 ACCTTGGCCCTTGTGTTGAGCAGTGGATGTGGAAGAGACACTGTTGTTTCAGCTACTGAG 3384

Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
   :::::::::::::::::::::
Db 3385 AGATTTTATGATCCCACTCAGTCGAGAAATTTGTTTGTATGATGATGATGCAAAAGACATA 3444

Qy 487 AsnValGlnTrpArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
```

```
Db 3445 AATATCCAGTGGCTGAGATCTTCACATCGGTATCTCTCACAAAGAGCCCAATCTGTTCAC 3504

Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
   :::::::::::::::::::::
Db 3505 TTCACCATCGCTGAAACATCGCTACGGGACAACAGCTCGGAGGTGTCCCATGAGNA 3564

Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlyGlyLeuProGluLys 546
   :::::::::::::::::::::
Db 3565 ATCATTAGTCAGCGAAAGCAGCCAGTATTCTATCTCTTCATTGACTCTCTGCCAGAAAA 3624

Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeu 566
   :::::::::::::::::::::
Db 3625 TACAATACTCGTGTAGGAGACAAGGGAACAGACTTCTTGGTGGTCAGAAAACAAGTATT 3684

Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
   :::::::::::::::::::::
Db 3685 GCTATTGCTCAGCTCTTATAGTAAAGCCAGATTCTGCTACTGGATGAAGCTACATCT 3744

Qy 587 AlaLeuAspAsnAspSerGlyValValGlnHisAlaLeuAspLysAlaArgThrGly 606
   :::::::::::::::::::::
Db 3745 GCCTTGGACACAGAAAGCGAAAGATTGTCCAGGAAGCACTGGATAAAGCCAGAGAAGGT 3804

Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
   :::::::::::::::::::::
Db 3805 CGCACCTGCATCGTAGACTCACCGCTGTGCCACCATCCAGAATGCTGCAAGATTGCT 3864

Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
   :::::::::::::::::::::
Db 3865 GTGATCCGAACGCAAGGCTCATGAGCAAGGACTCACAGAGCTCTCGTGGTGAAGAAAA 3924

Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
   :::::::::::::::::::::
Db 3925 GGCTTCTACTACTCGTGGTCAATGTTCAAAAGT 3957

RESULT 15
OAU78609 3858 bp mRNA linear MAM 11-OCT-2001
LOCUS Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
DEFINITION cds.
ACCESSION U78609
VERSION U78609.1 GI:2149086
KEYWORDS Ovis aries.
SOURCE Ovis aries
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
REFERENCE 1 (bases 1 to 3858)
AUTHORS Longley,M., Phua,S.H., van Stijn,T.C. and Crawford,A.M.
TITLE Isolation and mapping of the first ruminant multidrug resistance
genes
JOURNAL Anim. Genet. 30 (3), 207-210 (1999)
MEDLINE 99371931
PUBMED 10442984
REFERENCE 2 (bases 1 to 3858)
AUTHORS Longley,M. and Crawford,A.M.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1996) Biochemistry Department, University of
Otago, Dunedin, New Zealand
FEATURES
source
1. .3858
/organism="Ovis aries"
/db_xref="taxon:9940"
gene
1. .3858
/gene="mdr1"
1. .3858
/gene="mdr1"
/codon_start=1
/product="multidrug resistance protein-1"
/protein_id="AAB58489.1"
/db_xref="GI:2149087"
/translation="MDLEGDRNGRAAGNFKRDKRFFSKDKKRPVSTFTWTF"
```

RYNNWLDRLCMLVGLTAAIIHGAGLPLMVLFGDMTDSFAGAGNLTLSNLSNTST  
 IDRTYEGKFLKEMNTTAYIYIGAGVLIAIYQVSPFCLAGQVHRIRKQFPHAI  
 MQEIEGVDHVDHNTTNDVSKINEGIDKIMGFQPMFQWATFTTGIIFTGWN  
 LTVILIAIIPVLGSAIKWIKLSSFTDKLLAYAKAGAVEEVLAAIKTVIAFGQK  
 KLEIRYNKLEAKRIGIKKAITANISMGAAFLIYASALAFWFTGTSVLGREYSIG  
 QVLTVPFVLIGAFSIGOASPNTEAFANARGAAVEFKIIDNKPSITDSVNTGKRPDN  
 IKNLEFRNVHHPYPSRNEVKILKGLNLKXVSGQTVLVNCGCKSTTVQLMQRLYD  
 PTEGWSIDQDIRTINRYLREIIGVSOEPVLFAITTAENIRYRGREDVTWDEIQKA  
 VKEANAYDFIMKLPKFDLTVGERGAQLSGQKQRIATIALVRNPKLILLDEATSL  
 GTESEAVQALDKARKGRTTIVIAHRLSTVRNADIAGLDGVIVEGSHDELAKGR  
 GYFKLVTWTKGNELENTPEGLSKTDDLMSSODSRSLIRKSRPRIRGSQS  
 QDRKLTSTEDDESVPVFWFRLIKLINTWEPYVGVFCALINGALOPAFAPISRI  
 IGLFENDNDEYKRONSLFSLFLILIGLISITTFPLOGFTFGKAGEILTRRLRYLF  
 RSMLRQVSWFDPKNTGALTIRLANDAAQVKGAVSLAVITQNIANLIGLILISL  
 IYGMQUTLLULLAIPLIAVAGVIEKMLSGQALKKKLEAGKATEIAENFRVWS  
 LTRERFEYMYKQVPRYNSLRHAFVGTITSITQAMMYFSYACGRFPGALVAQ  
 IMEFQDLVYFSAVFGAMAVQVSTFADYAKAKVSAARVINIIEKPLIDSYSTEG  
 LKPSVTEGSAFNDVFNPTRPDPVPLRGLSLEVKGGOTLALVSSGCGKSTVQLL  
 ERFYDPLACTVFDCKEVKQNLVQMLRAHMGIVSQEPFLFDCSIENTAYGNSRVVS  
 QBEIEHAKEANIHSFEMLPDKYNTRVGDKGTQSGQKQRIATIALARLVROPHILL  
 DEATSALDTESEKVOEALDKAREGRTCVIAHRLSTIQNADLIYVFNQRIKEGHTH  
 QQLLAKQGIYFTMVSVQAGTKRQ"

BASE COUNT 1072 a 853 c 1002 g 931 t

## Alignment Scores:

Pred. No.: 2e-143 Length: 3858  
 Score: 1893.00 Matches: 364  
 Percent Similarity: 73.85% Conservative: 133  
 Best Local Similarity: 54.09% Mismatches: 158  
 Query Match: 56.97% Indels: 18  
 DB: 4 Gaps: 4

US-09-873-409-1 (1-659) x OAU78609 (1-3858)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
 Db 1828 GTCATCTGGAGGAGGACGATCATGATGAGCTCATGGGAAAAGAGGCATTTACTTCAA 1887  
 Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40  
 Db 1888 CTGTGTCACATGCG--ACAAAAGGAATGAATGGAATGGAAAATACCCCTGGTGAA 1944  
 Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60  
 Db 1945 TCCCTAAGTAAACTGATGACTTGTATCTCTCCCAAGATTCA---AGATCCAGTCTA 2001  
 Qy 61 Ile-----AspLysAlaGlu 65  
 Db 2002 ATCAGAGAAATCACTCGCAGAGTATCCGTGGATCCCAAGCCAGGACAGAAAGCTC 2061  
 Qy 66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysLysLys 85  
 Db 2062 AGTACAGAGAAACTTTGGATGAAGTGTACTCCCGTTTCTTTGGAGGATTCGTGAAG 2121  
 Qy 86 LeuAsnLysProGluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGly 105  
 Db 2122 CTGAATATAACTGAATGGCCCTATTTTGTGGTGTATTTTGGCCCATATAAATGGA 2181  
 Qy 106 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsn 125  
 Db 2182 GCTCTCAACACCATCTCAGTCATATTTCAAGGATTATAGGAATCTTCACCAAAAT 2241  
 Qy 126 AspLys---ThrThrLeuLysHisAspAlaGluIleIleIlePheValIleLeu 144  
 Db 2242 GACAATGACGAAACCAACGACAGACAGTAACTTCTTCTACCTGTGTCTTAATCCTT 2301  
 Qy 145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164  
 Db 2302 GGAATATTTCTTTTATTAATTTTCTTCCAGGGCTTCACGCTCGGCAAGCAGCGAG 2361  
 Qy 165 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 184  
 Db 184

Db 2362 ATCCTCACCAGGCGGCTCGGTACTCTGGTTTTCAGGTCGATGCTGAGACAGGTGTGAGC 2421  
 Qy 185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 204  
 Db 2422 TGGTTCGATGACCCCTAAACACACCCAGGAGCATTCGACGACGAGCTTGCCATGAGCC 2481  
 Qy 205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224  
 Db 2482 GCTCAAGTTAAAGGGGCTGTAGGTTCCAGGCTTGTGTAATATCCCAAGAATATGCAAT 2541  
 Qy 225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244  
 Db 2542 CTGGCAGCAGGATATATATATCCCTCATATGCTGCGAGTTAAACACTTCTCTCTGTTA 2601  
 Qy 245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264  
 Db 2602 GCAATTGTACCCATCATTCAGTAGCAGGAGTTATTGAAATGAAATGCTGCTGGACAA 2661  
 Qy 265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284  
 Db 2662 GCCTGAAAGATAAGAAAGAACTGGAAGGTGCTGGGAAGATTGCCCACTGAAGCAATAGAA 2721  
 Qy 285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304  
 Db 2722 AACTCCGAGCGTGTCTTCTTGTCTCGGAGGAGAGGTTGAAATATATATGATGCCCG 2781  
 Qy 305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324  
 Db 2782 AGTTTGCAGGTACCATACAGAACTCTTTCGAGGAGGCGCATGCTCTTTGGAATTACATTT 2841  
 Qy 325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344  
 Db 2842 TCCATCACTCAGGCAATGATGATTCTCTATGCTGGCTGTTTCCGGTTTGGTGCTAC 2901  
 Qy 345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364  
 Db 2902 TTGTGGCCCAAGGATCATGAGTTTCAGGATGTCTCTTAGTGTCTTCAGCTGTGTTC 2961  
 Qy 365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384  
 Db 2962 TTTGTGGCCATGGCTGTGGAGCAGGTCACTACATTTTGTCTCTGACTATGCCAAGGCCAA 3021  
 Qy 385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysLysProAsnIleAspSerArgSer 404  
 Db 3022 GTGTGACGACGCCGCTCATCAATATCATTTGAAAAAATTCCTCTGATTGACAGCTACAGC 3081  
 Qy 405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424  
 Db 3082 ACGGAAGGCTTAAGCCGAGTACAGTGGAAAGAGTGTGGCATTTAATGAGCTGTGTTC 3141  
 Qy 425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444  
 Db 3142 AACTACCCCACTCGACAGACGCTCCAGTGTCTCGGGGCTGAGCTGGAGGTGAAGAAG 3201  
 Qy 445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464  
 Db 3202 GCCCAGACGCTCGCGCTGTGGTGGAGTAGTGTGGTGGAGAGCAGCGCTGCTCCAGCTC 3261  
 Qy 465 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 484  
 Db 3262 CTGGAGCGGTTCTATGACCCCTTGGCTGGCAGCAGTGTATTATGAGCGCAAGGAATAAG 3321  
 Qy 485 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504  
 Db 3322 CAGCTGAATGTCCAGTGGCTGGGCGGACACATGGGCATCGTGTCCAGGAGGCCATTCTG 3381  
 Qy 505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524  
 Db 3382 TTTGACTGACGATCGCGGAGAAATATCGCTACGGAGACAAACAGCCGGGTGCTGTCCAG 3441  
 Qy 525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544  
 Db 3442 GAGGAGATTGACACCGGCCCAAGGAGGCCAAACATCCACTCTCTTATCGAGATGTGCGG 3501

Qy 545 GluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564  
Db 3502 GACAAATACAAACACCCAGAGTAGGGGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAG 3561  
Qy 565 ArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAspGluAla 584  
Db 3562 CGCATTGCCATAGCTCGCGCCCTTGTAGACAGCCTCACATCTTGCTTCTGGATGAAGCT 3621  
Qy 585 ThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604  
Db 3622 ACCTCGGCTCTGGACACCGAAAGTGAAAGGTGGTCCAAAGAGCCCTGGACAAAGCCCGA 3681  
Qy 605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624  
Db 3682 GAGGACGCACCTGCATCGTCATCGCTCACCGGCTGTCCACCATCCAGAACGCAGACTTG 3741  
Qy 625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArg 644  
Db 3742 ATAGTGTGTTCAGAAATGGCAGAAATCAAGGAGCATGGCAGACACACAGCAGCTGCTGGCG 3801  
Qy 645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657  
Db 3802 CAGAAAGGCATCTATTTCACCATGTGTCTCAGGCT 3840

Search completed: March 31, 2003, 02:07:24  
Job time : 8998.97 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 644.062 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3223

Sequence: 1 MLAEGAHAEIMAKRGLYY.....QELLNRNDIYFKLVNAQSVQ 659

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_epool/US09873409/runat\_27032003\_115416\_19216/app\_query.fasta\_1.7544  
-DB=N\_Geneseq\_101002 -Qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1.1 4247 @runat\_27032003\_115416\_19216 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3163	95.2	3699	24	Human transporter
2	1924.5	57.9	3924	21	Human ATP binding
3	1924.5	57.9	3924	21	Human MDR-3 DNA.
4	1924.5	57.9	3924	24	Gene #2299 used to
5	1882.5	56.7	3912	24	Rat sequence diffe
6	1882	56.6	4369	21	Rat multidrug resi
7	1882	56.6	4425	21	Rat multidrug resi
8	1873	56.4	4186	22	Cynomolgous monke
9	1873	56.4	4195	22	Cynomolgous monke
10	1872	56.3	2726	15	Multidrug-resistan
11	1872	56.3	2726	15	Multidrug-resistan
12	1872	56.3	2726	18	Multidrug-resistan
13	1872	56.3	3840	24	Multidrug resistan
14	1872	56.3	3860	21	Human polynucleoti
15	1872	56.3	3860	21	Human wild-type mu
16	1872	56.3	3860	24	Human BCRP DNA rel
17	1872	56.3	3860	24	Human BCRP DNA rel
18	1872	56.3	3988	21	Human MDR-1 DNA.
19	1872	56.3	4264	19	Mutated human P-gl
20	1872	56.3	4264	19	Mutated human P-gl
21	1872	56.3	4349	22	Human intestine ce
22	1872	56.3	4646	21	Human ATP binding
23	1872	56.3	4646	24	Human mdr1 Gene.
24	1872	56.3	4646	24	Kidney cancer rela
25	1872	56.3	4646	24	Kidney cancer rela
26	1872	56.3	4669	8	Sequence of human
27	1872	56.3	6505	17	Hybrid vector pSF-
28	1872	56.3	8630	21	Retroviral M4 mdr-
29	1872	56.3	8630	21	Retroviral vector
30	1871	56.3	4788	21	Murine multidrug r
31	1871	56.3	4788	21	Mouse BCRP DNA rel
32	1870	56.3	4279	22	Dog P-glycoprotein
33	1870	56.3	4279	22	Dog P-glycoprotein
34	1870	56.3	4279	22	Dog P-glycoprotein
35	1870	56.3	4279	22	Dog P-glycoprotein
36	1869	56.2	4279	22	Dog P-glycoprotein
37	1869	56.2	4378	11	Multidrug Resistan
38	1869	56.2	4669	14	Sequence of human
39	1861	56.0	4646	15	Human multidrug re
40	1855	55.8	4669	19	Human P glycoprote
41	1855	55.8	4669	24	cDNA encoding huma
42	1849.5	55.7	4233	21	Rat mdr1b2 (multis
43	1849.5	55.7	4233	22	Rat mdr1b2 multidr
44	1844.5	55.5	4189	21	Murine multidrug r
45	1844.5	55.5	4189	24	Mouse BCRP DNA rel

ALIGNMENTS

RESULT 1  
ABK83223

ID ABK83223 standard; cDNA; 3699 BP.

XX ABK83223;

AC ABK83223;

DT 27-AUG-2002 (first entry)

DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CBI, cDNA.

XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;  
neurological disorder; muscle disorder; immunological disorder; cancer;  
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;  
KW cell proliferative disorder; cervical cancer; breast cancer;  
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;  
KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;  
KW Grave's disease; gastrointestinal disorder; Crohn's disease;  
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;





Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360  
Db 2740 TTTGGAGCCTATTAAATCAAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTT 2799  
Qy 361 ThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380  
Db 2800 ACTGCAATTCATATGGAGCTATGGCCATCGGAAACGCCTGTTTTGGCTCCCTGAAATAT 2859  
Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400  
Db 2860 TCCAAAGCCAAATCGGGGCTGGCGATCTGTTGGCTTTGGTGAAGAAACCAATATA 2919  
Qy 401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420  
Db 2920 GACAGCCGAGTCAAGAAAGGCAAAAGCCAGACACATGTGAAGGGAAATTTAGAGTTTCGA 2979  
Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
Db 2980 GAAGTCTCTTTCTTATCCATGTGCGCCAGATGTTTCATCCTCCGGGCTTATCCCTC 3039  
Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
Db 3040 AGTATTGACGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACT 3099  
Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480  
Db 3100 TCTGTTCACTTCGCAGAGACTTATGACCCCGTGCAAGGACAGTCTGTTGATGGT 3159  
Qy 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500  
Db 3160 GTGGATGCAAAAGAAATTCGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCAA 3219  
Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
Db 3220 GAGCCTGCTCTTCACTGAGCAGTCTGAGAACATCGCTATGGTGGTGAACAGCCGT 3279  
Qy 521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540  
Db 3280 GTGGTGCCATTAGATGAGATCAAGAGCCGCAATGACGCAAAATATCCATTTTATT 3339  
Qy 541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560  
Db 3340 GAAGGTCTCCCTGAGAAATACACACAAAGTTGGACTGAAAGGAGCAGCAGCTTTCTGC 3399  
Qy 561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeu 580  
Db 3400 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAATTTTATTG 3459  
Qy 581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600  
Db 3460 TTGGATGAGCCACTTCAGCCCTCGATATGACAGTGAAGAGTGGTTCAGCATGCCCTT 3519  
Qy 601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620  
Db 3520 GATAAGCCAGGAGCGGAAGGACATGCTAGTGGTCACTCAGAGGCTCTCTGCAATTCAG 3579  
Qy 621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGluGlyThrHisGln 640  
Db 3580 AACCGAGATTTGATAGTGTCTGCACATGGAAGATTAAGAAACAAAGGAACACTCATCAA 3639  
Qy 641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659  
Db 3640 GAGCTCTCTGAGAAATCGACACATATATTTAAGTTAGTGAATGCACAGTCAGTCGAG 3696  
RESULT 2  
ID AA294742  
AA294742 standard; cDNA; 3924 BP.  
XX  
AC AA294742;  
XX  
DT 01-AUG-2000 (first entry)  
XX

DE Human ATP binding cassette ABCB4 (MDR3) cDNA.  
XX ABCB4; ATP binding cassette; human; cholesterol; lipid disorder;  
KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;  
KW lupus erythematosus; diagnosis; gene therapy; MDR3;  
KW multidrug resistance; chromosome 7q21; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200018912-A2.  
XX  
XX 06-APR-2000.  
XX  
XX 21-SEP-1999; 99WO-EP06991.  
XX  
XX 25-SEP-1998; 98US-0101706.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Schmitz G, Klucken J;  
XX  
XX WPI; 2000-293151/25.  
XX  
XX Adenosine triphosphate binding proteins useful for identifying agents  
PT for treating atherosclerosis and other inflammatory disorders -  
PT  
XX  
XX Claim 9; Page 117-119; 154pp; English.  
XX  
XX The present sequence is that of human ATP binding cassette  
CC subfamily B protein ABCB4 cDNA. The cDNA was identified using a  
CC differential display method in which monocytes from peripheral  
CC blood were subjected to macrophage differentiation and cholesterol  
CC loading with acetylated low density lipoproteins and subsequent  
CC delecting with high density lipoprotein (HDL3) to identify  
CC cholesterol sensitive genes. The gene maps to chromosome 7q21  
CC and is also termed MDR3 (multidrug resistance). The invention  
CC provides cholesterol-sensitive ABC genes (see AA294734-63). These  
CC genes, and polypeptides encoded by them, can be used for diagnostic  
CC and therapeutic applications, and for biochemical or cell-based  
CC assays to screen for pharmacologically active modulator compounds  
CC useful for the treatment of lipid disorders, atherosclerosis or  
CC other inflammatory diseases such as psoriasis and lupus  
CC erythematosus.  
XX  
SQ Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;  
Alignment Scores:  
Pred. No.: 5,43e-187 Length: 3924  
Score: 1924.50 Matches: 372  
Percent Similarity: 74.10% Conservative: 123  
Best Local Similarity: 55.69% Mismatches: 162  
Query Match: 57.91% Indels: 11  
DB: 21 Gaps: 3  
US-09-873-409-1 (1-659) x AA294742 (1-3924)  
Qy 1 MetLeuAlaGluLysGlyValaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 1851 GTAATTGGAGCAGGAGGAGCCACAGCCAACTGATGAGAGGAGGAGGGGTGTTACTTCAA 1910  
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31  
Db 1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAAGAAGATTTGAACATAATGAT 1970  
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51  
Db 1971 GAAAGGCTGCCACTAGAAATGCGCCCAATGCGTGAATCTCGCCTATTAGGCATCTCT 2030  
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69  
Db 2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAAGAGAGCCTTGATGTGGAACCGAT 2090  
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89

Db 2091 GCATTGAAGCAAAATGTCACCAGTGTCTTTCTGAAGGTCTCGAAATCGAATAAACA 2150  
QY 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisPro 109  
Db 2151 GAATGGCCCTACTTGTGTGGGAACAGTATGTCCTATGCCAATGGGGGCTTCAGCCG 2210  
QY 110 ValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129  
Db 2211 GCATTTTCAGTCATATCTCAGAGATCATAGCGATTTTGGACAGCGCATGTCAGTG 2270  
QY 130 LeuLysHisAaspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149  
Db 2271 AAGCAGCAGAAGTGCACATATTTCTCTTGAATTTCTTATTTCTGGGAATTTATTTCTTT 2330  
QY 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169  
Db 2331 TTTTACTTTCTCTTCCTTCAGGGTTTCAGTGTGGGAAGCTGGCGAGATCTCCACGAAGA 2390  
QY 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLys 189  
Db 2391 CTGGGTCAATGGCTTTTAAAGCAATGCTAAGACAGCATGAGCTGGTTTGATGACCAT 2450  
QY 190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209  
Db 2451 AAAAAACAGTACTGGTGCACTTTTACAAGACTTGCACAGATGCTGCCCAAGTCCAAGGA 2510  
QY 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229  
Db 2511 GCACAGGAACCAAGTGTGGCTTTAAATTTGACACAGAAATATAGCTTAACCTTGGAACTGGTAT 2570  
QY 230 IleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVal 249  
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAAACCTTATGCTATTAGCAGTTGTTCCAAT 2630  
QY 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269  
Db 2631 ATTGCTGTGTGAGGAATTTGTGAATTAAGAAATTTGCTGGCTGGAATGCCAAAGAGATAAA 2690  
QY 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289  
Db 2691 AAGAACTGGAAGCTGCTGGAAGAGTTGAACAGAGGCAATAGAAAATATTAGGACAGTT 2750  
QY 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309  
Db 2751 GTGCTTTGACCCAGGAAGAAATTTGAATCAATGATGTTGAAAAATTTGATGACCT 2810  
QY 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329  
Db 2811 TACAGGAATTTCTGCAGAAGGCACACATCTATGTAATTTACTTTTATGATCTCACAAAGCA 2870  
QY 330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349  
Db 2871 TTTATGATATTTTCTTCATCCCGTGTGTTTTCGATTTGGTGCATATCTCATTTGTAATGGA 2930  
QY 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleAlaTyrGlyAlaMetAla 369  
Db 2931 CATATCGGCTTCAGAGATGTTATTCGGTGTTTCTGCAATTTGTTATTTGGTCAGTGGCT 2990  
QY 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389  
Db 2991 CTAGGACATGCCAGTTTCATTTGCTCCAGACTATGCTAAAGCTAAGTGTCTGCAGCCAC 3050  
QY 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLys 409  
Db 3051 TTTATTCATCTGTTTGAAGACAAACCTCTGATTCAGAGCTACAGCTAAGAGGGGCTGAAG 3110  
QY 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429  
Db 3111 CTGATAAATTTGAAGAAATATAACATTTAATGAAGTGTGTTCAACTATCCACCCGA 3170  
QY 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449

Db 3171 GCAAAACGTGCCAGTCTTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGCCAGACACTAGCC 3230  
QY 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469  
Db 3231 CTGGTGGGACAGAGTGGCTGTGGGAAGACACGGTGGTCCAGCTCTGGAGCGGTCTTAC 3290  
QY 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489  
Db 3291 GACCCTTTGGGGGACAGTGTCTTCATGATGCTCAAGAAGCAAGAAATCAATGTCACAG 3350  
QY 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509  
Db 3351 TGGCTCAGAGCTCAATCGGAATCGTCTCAGGAGCCTATCTTATTTGACTGACGCAAT 3410  
QY 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529  
Db 3411 GCGGAGAATATTGCTATGGAGAACACCGGGTGTATCACAGATGAATTTGTGAGT 3470  
QY 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549  
Db 3471 GCAGCCAAAGCTGCCAACATACATCTTTTCATCGAGACGTTACCCCAACAATATGAACA 3530  
QY 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569  
Db 3531 AGAGTGGGAGATGAAGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590  
QY 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAsp 589  
Db 3591 CGAGCCCTCATCAGACCAACCTCAAATCTCTCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650  
QY 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609  
Db 3651 ACTGAAAGTGAAGAGTTGTCCAAGAGCCCTGGCAAAAGCCAGAGAAGCCGCACTGC 3710  
QY 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629  
Db 3711 ATTTGATTTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAG 3770  
QY 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649  
Db 3771 AATGGAGAGCTCAGAGCATGGCAGCATGGCAGCTCAGAGCTTAAATAGTGTGTTTCAG 3830  
QY 650 PheLysLeuValAsnAlaGlnSer 657  
Db 3831 TTTTCAATGCTCAGTGTCCAGCT 3854  
RESULT 3  
AAZ88974  
ID AAZ88974 standard; DNA; 3924 BP.  
XX AAZ88974;  
AC AAZ88974;  
XX  
DT 30-MAY-2000 (first entry)  
XX Human MDR-3 DNA.  
DE  
XX Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;  
KW stomach; colon; ulcerative colitis; pouchitis; large intestine; human;  
KW inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;  
KW ileum; anti-ulcer; anti-inflammatory; cytostatic; ss.  
OS Homo sapiens.  
XX  
PN WO200007577-A2.  
XX  
PD 17-FEB-2000.  
XX  
PF 06-AUG-1999; 99WO-DE02426.  
XX  
PR 06-AUG-1998; 98DE-1035526.  
PR 15-DEC-1998; 98DE-1057750.  
XX  
PA (STRE/) STREMMEL W.

XX  
PI Stremmel W;  
XX  
XX WPI; 2000-195439/17.  
XX  
XX Using phosphatidylcholine having mucous membrane protective activity to  
PT prevent or treat large intestinal diseases, such as ulcerative colitis  
PT or pouchitis -  
XX  
XX Disclosure; Page 17-22; 22pp; German.  
XX  
XX This invention describes a novel medicament containing a therapeutically  
CC effective amount of phosphatidylcholine to treat diseases, where  
CC phosphatidylcholine has an advantageous mucous membrane protective  
CC activity in the large intestine. The invention describes (1) a method  
CC to regulate presence or absence of MDR (multidrug resistance)-3  
CC transfection in a sample from patients with suspicion of stomach and  
CC colon diseases, where treatment with phosphatidylcholine is indicated;  
CC (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large  
CC intestinal inflammation, Crohn's disease, diverse colitis, infectious  
CC enteritis/colitis, inflammation through X-ray treatment, antibiotics,  
CC chemotherapeutics, drugs or chemicals or large intestinal carcinoma,  
CC contains MDR3-specific primers to detect mutations in the MDR3 gene  
CC family and their interactions partners (in particular transcription  
CC factors); and (3) a method to manufacture a medicament for local gene  
CC therapy in the ileum with the ideal MDR3-analogous genes for patients  
CC with missing or decreased expression of MDR3-analogous proteins. The  
CC products of the invention have anti-ulcer, anti-inflammatory and  
CC cytostatic activity. The orally administered delayed-release form of  
CC phosphatidylcholine prevents premature resorption and provides for  
CC targeted release in the lower section of the small or large intestine.  
CC This sequence encodes the human MDR-3 protein described in the method of  
CC the invention.  
XX  
XX Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;  
XX

Alignment Scores:  
Pred. No.: 5,43e-187 Length: 3924  
Score: 1924.50 Matches: 372  
Percent Similarity: 74.10% Conservative: 123  
Best Local Similarity: 55.69% Mismatches: 162  
Query Match: 57.91% Indels: 11  
DB: 21 Gaps: 3

US-09-873-409-1 (1-659) x AAZ88974 (1-3924)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 1851 GTAAATTGGGAGCAAGGAAGGCACAGCGAACTGTATGAAGAAGAGGGGTGTACTCAA 1910  
Qy 21 LeuVal-----MetSerGlnAspIleLysLysAlaAsp 31  
Db 1911 CTTGTCAACATGCGACACATCAGGAAGCCAGATCAGTCAGAGAAATTGAACTAAATGAT 1970  
Qy 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51  
Db 1971 GAAAAGGCGCCACTAGATGGCCCCCAATGGCTGGAAATCTCGCCCTATTAGGCATTCT 2030  
Qy 52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69  
Db 2031 ACTCAGAAAAACCTTAAAAAATTCACAAATGTGTGAGAGAGCCCTTGATGGAAACCGAT 2090  
Qy 70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89  
Db 2091 GGACTTGAAGCAAAATGTCACACAGTGCTCTTCTGAAGGTCCTGAAACCTGAATAAACA 2150  
Qy 90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109  
Db 2151 GAATGGCCCTACTTGTGCTGGGAACAGTAGTGCCCATGTGCCAATGGGGGCTTCAGCG 2210  
Qy 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129  
Db 2211 GCATTTTCAGTCATATCTTCAGAGATCATACGCAATTTTGGACCAGCGCATGTGCAGTG 2270

QY 490 TTPLeuArgSerGlnIleAlaIleValProGlnGluProValIleuPheAsnCysSerIle 509  
 Db 3351 TGCTCAGAGCTCAACTCGGAATCGTGCTCAGAGCGCTATCTATTGACTCGACGATT 3410  
 QY 510 AlaGluAsnIleAlaIleValProGlnGluProValIleuPheAsnCysSerIle 529  
 Db 3411 GCCGAGAATATTCCTATGGAGACACAGCGCGTGTATCAGAGATGAAATGTGAGT 3470  
 QY 530 AlaAlaAsnAlaAlaIleHisSerPheIleGluGluLeuProGluIleValIleuPhe 549  
 Db 3471 GCAGCCAAAGCTGCCAACATACATCTCTTCATCGAGACGTTACCCACAAATATGAACA 3530  
 QY 550 GlnValClyLeuIleGlyAlaGlnLeuSerGlyGlyGlnIleGlnArgLeuAlaIleAla 569  
 Db 3531 AGAGTGGGAGATGAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATGTCTATTGCC 3590  
 QY 570 ArgAlaLeuLeuGlnIleProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589  
 Db 3591 CGAGCCCTCATCAGACAACTCAATCTCTCTGTGTGATGAGCTACATCAGCTCTGGAT 3650  
 QY 590 AsnAspSerGluIleValValGlnHisAlaLeuAspLysAlaIleValIleuPheThrCys 609  
 Db 3651 ACTGAAAGTGAAGGTTGTCCAAAGAGCCCTCGACAAAGCCAGAGAGCGCACCTGC 3710  
 QY 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629  
 Db 3711 ATGTGATGTCTCAGCGCTGTCCACATCCAGATCCAGATTAATAGTGTGTTCAG 3770  
 QY 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyr 649  
 Db 3771 AATGGAGAGTCAAGGAGCATGCGACGATCAGCAGTCTGTCGACAGAAGGCATCTAT 3830  
 QY 650 PheLysLeuValAsnAlaGlnSer 657  
 Db 3831 TTTTCAATGTCAGTGTCCAGGCT 3854  
 RESULT 4  
 ABN95801  
 ID ABN95801 standard; DNA; 3924 BP.  
 XX AC ABN95801;  
 XX DT 13-AUG-2002 (first entry)  
 XX DE Gene #2299 used to diagnose liver cancer.  
 XX KW Gene; liver cancer; da; hepatocellular carcinoma; hepatotrophic;  
 KW metastatic liver tumor; cytostatic; expression profile; disease state;  
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX OS Homo sapiens.  
 XX PN WO200229103-A2.  
 XX PD 11-APR-2002.  
 XX PF 02-OCT-2001; 2001WO-US30589.  
 XX PR 02-OCT-2000; 2000US-237054P.  
 XX PA (GENE-) GENE LOGIC INC.  
 XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
 XX WPI; 2002-426119/45.  
 XX PT Diagnosing and detecting the progression of liver cancer,  
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
 PT involves detecting the level of expression of two or more genes in a  
 PT liver tissue sample -  
 XX PS Claim 1; SEQ ID NO 2299; 298pp; English.

CC The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumour in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5,43e-187 Length: 3924  
 Score: 1924.50 Matches: 372  
 Percent Similarity: 74.10% Conservative: 123  
 Best Local Similarity: 55.69% Mismatches: 162  
 Query Match: 57.91% Indels: 11  
 Ds: 24 Gaps: 3  
 US-09-873-409-1 (1-659) x ABN95801 (1-3924)  
 QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
 Db 1851 GTAATGTGGACAGAGGAGCCACAGCGNACTGATGAAGAGGAGGGGTGACTTCAA 1910  
 QY 21 LeuVal-----MetSerGlnAspIleLysLysLysAla 31  
 Db 1911 CTTGTCAACATGCAGACATCAGAGCCAGAGCCAGATCCAGTCAAGAAATTTGAACATAATGAT 1970  
 QY 32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51  
 Db 1971 GAAAAGGCTGCCCATAGAAATGCCCAAAATGGCTGCGAAATCTCGCCTATTTAGGCATTCT 2030  
 QY 52 SerValLysSerIleLys--SerAspPheIleAspLysAla--GluGluSerThrGln 69  
 Db 2031 ACTCAGAAAACCTTAAATAATTCACAAATGTGTACAGAGAGCCTTGATGTGAAACCGAT 2090  
 QY 70 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 89  
 Db 2091 GGACTTGAAGCAAAATGTCACCAGTGTCTTCTGGAAGTCTCTGAAACTGAAATAAACA 2150  
 QY 90 GluTyrProPheValIleuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109  
 Db 2151 GAATGGCCCTACTTTGTCTGGGAACAGTATGTGCCATTTGCCAATGGGGGGCTTCAGCG 2210  
 QY 110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129  
 Db 2211 GCATTTTCAGTCATATTTCTCAGAGATCATAGCGATTTTGGACACGCGATGATGACAGTG 2270  
 QY 130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149  
 Db 2271 AAGCAGAGAAGTGCAACATATCTCTTGATTTCTTATTCTTCTGGAATTAITCTTTT 2330  
 QY 150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169  
 Db 2331 TTTACTTTCTCTTCAGGGTTTTCAGTTTGGGAAAGCTGGCGAGATCTCTCCAGAGA 2390  
 QY 170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLys 189  
 Db 2391 CTGGCTCATGCTTTTAAACAAATGCTAAGACAGACATGAGCTGTTTGTATGACCAT 2450  
 QY 190 GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209  
 Db 2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510  
 QY 210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229

Db 2511 GCCCAGGACACGAGTGGCTTTAATGTCACAGAAATATAGCTTAACCTTGGAACTGGTATT 2570  
Qy 230 lleileSerPheileTyrglyTyrGluMetThrPheLeuileLeuSerileAlaProVal 249  
Db 2571 ATCATATCATTTATCTACGGTTGGCAGTTAAACCTATTGCTATTAGCAGTTGTTCCAAAT 2630  
Qy 250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269  
Db 2631 ATTGCTGTGTACAGGAATTTGTTGAAATGAAATTTGGTGGGAAATGCCCAAGAGATATA 2690  
Qy 270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289  
Db 2691 AAAGAACTGGAAAGCTGCTGGAAGATTTGCAACAGAGCCCAATAGAAAATATTAGGACAGTT 2750  
Qy 290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLysGlnThrGln 309  
Db 2751 GTGTCTTTGACCAGCAAGAAATTTGAAATCAATGATGTTGAAATTTGATGGACCT 2810  
Qy 310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329  
Db 2811 TACAGGAATTTCTGTGAGAGGCACACATCTATGGAATTTACTTTTAGTATCTCACAAGCA 2870  
Qy 330 PheIleTyPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349  
Db 2871 TTTATGTAATTTCTATGCTGCGGTTGTTTCGATTTGGTGCAATATCTCATGTTGAATGGA 2930  
Qy 350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369  
Db 2931 CATATGCGCTTCAGAGATGTTATCTGTGTTTCTGCNAATGTTATTTGGTCAGTGCT 2990  
Qy 370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389  
Db 2991 CTAGGACATGCCAGTTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGCTCAGCCCCAC 3050  
Qy 390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409  
Db 3051 TTTATTCATGCTGCTTTGAAAGACAACCTCTGATTGACAGTCAAGTGAAGAGGGGCTGAAG 3110  
Qy 410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429  
Db 3111 CTGTATAAATTTGAAGAAATATACATTTAATGAAGTCGTGTTCAACTATCCACCGCA 3170  
Qy 430 ProAspValPheIleLeuArgGlyLeuSerLeuSerileGluArgGlyLysThrValAla 449  
Db 3171 GCAAAAGTGCAGTCTTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230  
Qy 450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469  
Db 3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGACACGGTGTCCAGCTCTCTGGAGCGTTCTAC 3290  
Qy 470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489  
Db 3291 GACCCCTTGGCGGGGACAGTGTCTTCGATGTGCAAGAAACAAAGAACTCAATGTCCAG 3350  
Qy 490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509  
Db 3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGAGCCTATCTCTATTGACTGCAGCAT 3410  
Qy 510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529  
Db 3411 GCGGAGAATATTGCGCTATGGAGACAACACGCGGTTGTATCACAGGATGAAATTTGTGAGT 3470  
Qy 530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549  
Db 3471 GCAGCAAGCTGCCAACATACATCTTTTCATCGAGACGTTACCCACAAATATGAACA 3530  
Qy 550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569  
Db 3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGAGGTCAAAACAGAGGATTCCTATTGCC 3590  
Qy 570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589

Db 3591 CGAGCCCTCATCAGACAACCTCAAAATCTCTCTGTTGATGAAGCTACATCAGCTCTGGAT 3650  
Qy 590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609  
Db 3651 ACTGAAAGTGAAGAGTTGTCCAAGAGCCCTGGACAAGCCAGAGAAGCGCGCACCTGC 3710  
Qy 610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629  
Db 3711 ATTGTGATTTGCTACCGCCTGTCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAG 3770  
Qy 630 AsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIleTyr 649  
Db 3771 AATGGAGAGTCAGAGCATGGCAGCATCAGCAGCTGCTGGCACAAAGGCATCTAT 3830  
Qy 650 PheLysLeuValAsnAlaGlnSer 657  
Db 3831 TTTTCAATGGTCAGTGTCCAGGCT 3854  
RESULT 5  
ABK63653  
ID ABK63653 standard; cDNA; 3912 BP.  
XX  
AC ABK63653;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Rat sequence differentially expressed in response to a hepatotoxin #1560.  
XX  
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
KW differential expression; centrilobular necrosis; steatosis.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200210453-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 30-JUL-2001; 2001WO-US23872.  
XX  
PR 31-JUL-2000; 2000US-222040P.  
PR 02-NOV-2000; 2000US-244880P.  
PR 11-MAY-2001; 2001US-290029P.  
PR 15-MAY-2001; 2001US-290645P.  
PR 22-MAY-2001; 2001US-292336P.  
PR 06-JUN-2001; 2001US-295798P.  
PR 13-JUN-2001; 2001US-297457P.  
PR 19-JUN-2001; 2001US-298884P.  
PR 09-JUL-2001; 2001US-303459P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Mendrick D, Porter MW, Johnson KR, Castile AL, Elashoff MR;  
XX  
XX WPI; 2002-241625/29.  
XX  
PT Predicting toxic effects of compounds or the progression of these toxic  
PT effects by determining the changes in gene expression in tissues or  
PT cells exposed to the toxin and comparing these to gene expression in  
PT unexposed tissues or cells -  
XX  
PS Claim 1; Seq ID No 1560; 239pp; English.  
XX  
CC The invention relates to methods for predicting toxic effects of  
CC compounds or the progression of these toxic effects by determining the  
CC global changes in gene expression in tissues or cells exposed to the  
CC toxin and comparing these to gene expression in unexposed tissues or  
CC cells. Also included are methods of predicting at least one toxic  
CC effect of a compound or progression of a toxic effect, preferably the  
CC hepatotoxicity of a compound, comprising detecting the level of  
CC expression in a tissue or cell sample exposed to the compound of two or  
CC more genes listed in the specification, where differential expression of  
CC the genes is indicative of at least one toxic effect or progression.  
CC The method can also be used to identify an agent which modulates the





Ds 3522 GAACAAGAGTAGGACAGAGGGGACAGAGCTCTCTGGAGGCCAGAAACAGAGGATTGCT 3581  
Qy 568 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 587  
Dd 3582 ATCGCCCGAGCCCTCATCAGACAGCTCGGGTCTACTGCTGGATGAAGCCACAGTCGGCT 3641  
Qy 588 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuLeuAspLysAlaArgThrGlyArg 607  
Dd 3642 TTGGACACTGAGAGTGAAAGGCTCTCCAGGAAGCGCTGGACAAAGCCAGGAGGCGCGC 3701  
Qy 608 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 627  
Dd 3702 ACCTGCATGTGATCGGCACCGCTGTCACCATCCAGAGCCAGACTTGTATGCTGGTG 3761  
Qy 628 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 647  
Dd 3762 ATCGACAACGGCAAGGCTCAAGGAGCAGCGCACCCACAGCAGCTGCTGGCCCGAAGAGC 3821  
Qy 648 IleTyrPheLysLeuValAsnAlaGlnSer 657  
Dd 3822 ATCTATTCTTCATGCTCAACATTCAGAGCT 3851  
RESULT 6  
AAZ52047  
ID AAZ52047 standard; cDNA; 4369 BP.  
XX AC AAZ52047;  
XX AC  
XX DT 18-JUL-2000 (first entry)  
XX DE Rat multidrug resistance protein la cDNA.  
XX KW Multidrug resistance protein la; mdrla; multi-specific drug transporter;  
KW drug formulation; formulation excipient; compound design; inflammation;  
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;  
KW central nervous system disorder; auto-immune disease; kidney disease; ss.  
XX OS Rattus rattus.  
XX FH Key Location/Qualifiers  
XX FT CDS 352..4170  
XX FT /\*tag= a  
XX FT /product= "Rat multidrug resistance protein"  
XX PN WO200015650-A1.  
XX PD 23-MAR-2000.  
XX PF 10-SEP-1999; 99WO-US20770.  
XX PR 17-SEP-1998; 98US-0156800.  
XX PR 09-DEC-1998; 98US-0208809.  
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.  
XX PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;  
XX WPI; 2000-271372/23.  
XX DR P-PSDB; AAY70596.  
XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays  
XX to provide information on drug formulation, selection of formulation  
XX excipients and compound design -  
XX Claim 2; Page 23-24; 33pp; English.  
XX The present cDNA sequence encodes rat multidrug resistance  
XX protein la (mdrla). This protein is a member of the multi-specific drug  
XX transporters family. Mdr1a is used in assays to provide information on  
XX drug formulation, selection of formulation excipients and compound  
XX design. They are used in cell based, membrane based, binding or other  
XX assays to provide information that may enhance drug formulation. This  
XX invention further relates to the generation of in vivo and in vitro

CC comparison data to predict oral absorption and pharmacokinetics. This  
CC enables the selection of drugs with optimal pharmacokinetics, i.e. good  
CC oral bioavailability, brain penetration, plasma half life, and minimum  
CC drug interaction. Transgenic and knock-out animals created using rat  
CC mdrla provides an insight into treating and preventing human diseases  
CC including cancer, inflammation, cardiovascular disease, central nervous  
CC system disorders, auto-immune and kidney disease.  
XX SQ Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

## Alignment Scores:

Pred. No.:	1,49e-182	Length:	4369
Score:	1882.00	Matches:	362
Percent Similarity:	74.37%	Conservative:	137
Best Local Similarity:	53.95%	Mismatches:	158
Query Match:	56.64%	Indels:	14
DB:	21	Gaps:	5

US-09-873-409-1 (1-659) x AAZ52047 (1-4369)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Dd 2140 GTCATTGTGGAGCAAGAAATCATGCTCATGAGCTCATGAGAGAGAAAGAAATTTACTTCAA 2199  
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer 36  
Dd 2200 CTTGTCTCATGCTCAGACAGCAGGAAATGAAATTTAGGAATGAAAGCTTGTGAATCT 2259  
Qy 37 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu----- 48  
Dd 2260 AAAGACGGAAATTGATATGTGGACATGCTTCAAAAGAAATTCGGGATCCAGCTAATAGA 2319  
Qy 49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68  
Dd 2320 AGAAGATCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAATTTAGCACC 2379  
Qy 69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
Dd 2380 AAAGAGGCTCTGGATGACGACGCTACCTCCAGCTCTCTTTGGCGGATCTCTGAAGTTGAAT 2439  
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
Dd 2440 TCAACTGAATGGCTTATTTGTGGTGGTGTATTTGTGGCATTAATAATGGAGGCTTG 2499  
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127  
Dd 2500 CAACAGCATCTCTCATTAATATTTCAAGGTGTAGGGGTTTTACAAAATATGCACC 2559  
Qy 128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146  
Dd 2560 CCTGAAATCCAGCGCAGACAGCAACTTGTCTTTTATTGTTCTCTGATCTCTGGATC 2619  
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166  
Dd 2620 ATCTCTTTTACGTTTTTCTCTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCTCT 2679  
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186  
Dd 2680 ACCAAGCGACTCCGATACATGGTCTTCAATCCATGCTGAGACAGCATAGCTGGTTT 2739  
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
Dd 2740 GATGACCTTAAACACACACAGGCGCTGACCACAGGCTTGCCTCAATGACGCTCTCAA 2799  
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
Dd 2800 GTGAAGGGGCTACAGGCTCTAGGCTCTGCTGTATTATCCAGAAATAGCAATCTGGG 2859  
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleSerIle 246  
Dd 2860 ACAGGCATCATATCCCTCATCTACGCTGGCAATTTGACACTTTTCTCTAGCAATT 2919  
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266



```
Db 2920 GTTCCCATCTTCTATAGCAGGAGTGTGAAATGAAATGTTGTGACAAAGCGCTG 2979
Qy 267 LybAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2980 AAGATAAGAGGAACCTAGAGGTTCTGGAGACATCGCTACAGAGCAATTGAACACTT 3039
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 3040 CGCACTGTGCTCTTGTGACTCGGAGCAGAACTTGAACATATGATGATGCCAGACTTG 3099
Qy 307 GlnThrGlnHisArgSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3100 CAGATCCATACAGAAATGCTTTGAGAGAGCGCAGCTCTTGGGATCACTTCTCTCTTC 3159
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db 3160 ACCCAGGCCATGATGATTTCTCTATGCTGTTGTTCCGGTTGATGCTACTTGGTG 3219
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3220 GCACGAGAACTCATGACATTTGAAATGTTCTGTAGTATTTCTCAGTATTGCTTTGGT 3279
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3280 GCATCGCAGTGGGCGAGTCACTTCACTCGCTCTGACTACCGGAAAGCCAAAGTCTCG 3339
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3340 GCATCCCATCATCATCAGGATCATTTGAGAAATCCCTGAGATTGACAGCTACACGCGAG 3399
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db 3400 GGCTTCAAGCCTAATATGTTGGAAGCAAAATGTCAAAATTAATGGAGTCATGTTCAACTAT 3459
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3460 CCCACCAGCCCAACATCCAGTGTCTCAGGGGCTGAGCTAGAGTGAAGAAAGGCGAC 3519
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 3520 ACCTGGCCCTCTGGGCGCAGTGGCTGGGGAAGATACAGTGTCTCAGCTGCTTGAG 3579
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3580 CGCTTCTATGACCCCATCGCGCCACAGTGTCTTAGATGGCAAGAAATAAAGCAACTC 3639
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnLeuProValLeuPheAsn 506
Db 3640 AATGTCCAGTGGCTCGCGCCACCTGGGCATTTGTCTCCAGGAGCCCATCTCTGTTGAC 3699
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3700 TGCAAGATCGCCGAGAACATGTCTCAGGAGACACAGCGGTCTGTCTCATAGGAG 3759
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3760 ATCGTGAAGGCCCAAGAGGCGCAACATCCACAGTTTCATCGACTCACTGCTGAGAAA 3819
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3820 TACAACACCAAGTGGGAGACAAAGGACTCACTGTCTCGGGCGGAGAGCAGCGCATC 3879
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3880 GCATCGCGCGCCCTCTGTCAGACAGCTCATCTTACTTCTGATGAGGACATCA 3939
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3940 GCTCTGATACGGAGAGTGAAGAGTCTGTCAGAGCGCTGGACAAAGCCAGGGAAGGC 3999
Qy 607 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
```

```
Db 4000 CGCACTGCATTGTGATCGCGCACCGCCTGTCCACATCCAGACGACACTTGATCGTG 4059
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgSerArg 646
Db 4060 GTGATTTCAGAACGGCAGGTCAAGGACGACGCGCACCCACGACGCTGTGGCCAGAAA 4119
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 4120 GGCATCTATTTCGATGCTGTCAGTGTGCGAGCT 4152

RESULT 7
AAZ52048
ID AAZ52048 standard; cDNA; 4425 BP.
XX
AC AAZ52048;
DT 18-JUL-2000 (first entry)
XX
DE Rat multidrug resistance protein la cDNA derived from EST sequences.
XX
KW Multidrug resistance protein la; mdrla; multi-specific drug transporter;
KW drug formulation; formulation excipient; compound design; inflammation;
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW central nervous system disorder; auto-immune disease; kidney disease;
KW EST; expressed sequence tag; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT CDS 352..4170
FT /*tag= a
FT /product= "Rat multidrug resistance protein"
XX
PN WO200015650-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1999; 99WO-US20770.
XX
PR 17-SEP-1998; 98US-0156800.
PR 09-DEC-1998; 98US-0208809.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
XX
DR WPI: 2000-271372/23.
XX
DR P-FSDB; AAY70597.
XX
PT Isolated rat mdrla polynucleotides and polypeptides, useful in assays
PT to provide information on drug formulation, selection of formulation
PT excipients and compound design -
XX
PS Claim 11; Page 27-29; 33pp; English.
XX
CC The present cDNA sequence encodes rat multidrug resistance protein la
CC (mdrla). This cDNA is derived from EST (expressed sequence tag)
CC sequences. Mdrla is used in assays to provide information on
CC drug formulation, selection of formulation excipients and compound
CC design. They are used in cell based, membrane based, binding or other
CC assays to provide information that may enhance drug formulation. This
CC invention further relates to the generation of in vivo and in vitro
CC comparison data to predict oral absorption and pharmacokinetics. This
CC enables the selection of drugs with optimal pharmacokinetics. i.e. good
CC oral bioavailability, brain penetration, plasma half life, and minimum
CC drug interaction. Transgenic and knock-out animals created using rat
CC mdrla provides an insight into treating and preventing human diseases
CC including cancer, inflammation, cardiovascular disease, central nervous
CC system disorders, auto-immune and kidney disease.
XX
SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:
```

Pred. No.:	1,528-182	Length:	4425
Score:	1982.00	Matches:	362
Percent Similarity:	74.3%	Conservative:	137
Best Local Similarity:	53.9%	Mismatches:	158
Query Match:	56.64%	Indels:	14
DB:	21	Gaps:	5

US-09-873-409-1 (1-659) x AAZ52048 (1-4425)

Pred. No.:	1.52e-182	Length:	4425
Score:	1882.00	Matches:	362
Percent Similarity:	74.37%	Conservative:	137
Best Local Similarity:	53.95%	Mismatches:	158
Query Match:	56.64%	Indels:	14
DB:	21	Gaps:	5

US-09-873-409-1 (1-659) x AAZ52048 (1-4425)

Qy	1	MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	2140	GTCATTGTGGACCAAGGAATCATGATGAGCTCATGAGAGAGAAGGAATTTACTTC	2199
Qy	21	LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer	36
Db	2200	CTTGTGCATGACTCAGACAGCAGGAATAAATGAAATAGGAATGAAGCAAGCTTGTGAATCT	2259
Qy	37	-----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----	48
Db	2260	AAAGAUGGAATTGAAATGTGGACATGCTCTTCAAAAGAAATTCGGATCCAGTCTAATAAGA	2319
Qy	49	ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr	68
Db	2320	AGAAGATCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACCTTAGCACC	2379
Qy	69	---GlnSerLysGluLysSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	87
Db	2380	AAAGAGGCTCTGGATGACGACATCCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAAT	2439
Qy	88	LysProGluTyrProPheValValLysGlyThrLeuAlaSerValLeuAsnGlyThrVal	107
Db	2440	TCAACTGAATGGCCCTATTTTGTGGTGTGTATTTTGTGCCATAATAATGGAGCCTTG	2499
Qy	108	HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys	127
Db	2500	CAACCCAGCATTCCTCCATAATAATTTCAAAAGTGTAGGGGTTTTCACAAAAAATGACACC	2559
Qy	128	ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	146
Db	2560	CTGAAATCCACGGCAGAACAGCAACTTGTCTTTTATTTATTTCTGATCTCTGGGATC	2619
Qy	147	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	166
Db	2620	ATCTCTTTTCATTACGTTTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGATCCTC	2679
Qy	167	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe	186
Db	2680	ACCAACGAGCTCCGATACATGCTTCAAAATCCATCTGCTGAGACAGGACATAAGCTGTTT	2739
Qy	187	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	206
Db	2740	GATGACCCCTAAMAAACACACAGAGCGCTGACACACAGCTTGCCCAATGACGCTGCTCAA	2799
Qy	207	IleGlnGlyValaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	226
Db	2800	GTGAAGGGGCTACAGGCTTAGGCTGTGTTATTTACCCAGNACATAGCAAAATCTTGGG	2859
Qy	227	LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle	246
Db	2860	ACAGGCATCATCATCTCCCTGATCTACGGCTGGCAATTGACACTTTTACTCTCTAGCAATT	2919
Qy	247	AlaProValLeuAlaValThrGlyMetIleIleGluThrAlaAlaMetThrGlyPheAlaAsn	266
Db	2920	GTTCCCATCATTGCTATACAGAGGTGGTTGAATGAAATGTTGTCTGGACAAAGCGCTG	2979
Qy	267	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	286
Db	2980	AAAGATAAGAAGAACTAGAAGTTCGGAAGATCGCTACAGAAGCAATTGAAACTTTT	3039
Qy	287	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	306
Db	3040	CGCACTGCTCTTTGATCTGGGACAGAGAAGTTTGAACATATGTATGCCACAGACCTTG	3099

AAF86127  
 ID AAF86127 standard; cDNA; 4186 BP.  
 XX  
 AC AAF86127;  
 XX  
 DT 25-JUN-2001 (first entry)  
 XX  
 DE Cynomologous monkey P-glycoprotein cDNA.  
 XX  
 KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;  
 KW efflux pump; ss.  
 XX  
 OS Macaca fascicularis.  
 XX  
 FH Key Location/Qualifiers  
 CDS 100..3942  
 FT /\*tag= a  
 FT /product= "PGP"  
 FT /note= "P-glycoprotein"

XX WO200123565-A1.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26592.

XX 28-SEP-1999; 99US-0156921.

XX 12-OCT-1999; 99US-0158818.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steimel-Crespi DT, Crespi CL;

XX WPI; 2001-316136/33.

XX P-PSDB; AAB81064.

XX Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 (PGP) and homologous PGP polypeptides are useful for predicting  
 bioavailability of compound and increasing PGP transporter activity in  
 cell

XX Example 1; Page 51-57; 84pp; English.

CC This invention relates to a polynucleotide sequence encoding a  
 cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
 of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 exporting small molecules across the cell membrane. The invention  
 includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 and protein, and also that of an allelic variant. The PGP polynucleotide  
 sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 activity in a mammalian cell. They may also be used for increasing the  
 bioavailability of a drug. The present sequence represents cDNA encoding  
 CC cynomologous monkey P-glycoprotein.

XX Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:

Pred. No.:	1-178-181	Length:	4186
Score:	1873.00	Matches:	354
Percent Similarity:	74.23%	Conservative:	146
Best Local Similarity:	52.60%	Mismatches:	155
Query Match:	56.36%	Indels:	18
DB:	22	Gaps:	5

US-09-873-409-1 (1-659) x AAF86127 (1-4186)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20

DB 1912 GTCAATTGTGAGAAAGAAATCATGATGAGCTCATGAAGAGAAAGCAATTTACTTCAA 1971

QY 21 LeuValMetSerGlnAspIle-----LysLysAlaAspGlu--- 32

Db	1972	CTTGTCACAAATCCAGACAGCAGGAAATGAATTAATGAAATAGAAAATGCAGCTGATGAATCC	2031
QY	33	-----GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----	48
DB	2032	AAAAGTGAATTTGATACCTTGGAAATGCTTCACATGATTCAGGATCCAGTCTAATAAGA	2091
QY	49	-----ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGlu	65
DB	2092	AAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA-----GACAGAAAGCTT	2145
QY	66	GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysLysLysLysLys	85
DB	2146	AGTACCAAGAGGCTCTGGATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTAAGAAG	2205
QY	86	LeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly	105
DB	2206	CTAAATTTAACTGAGTGGCTTATTTTGTGTGGTGATTTTGTGCCATTAATAATGA	2265
QY	106	ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsn	124
DB	2266	GGTCTGCAACACAGCATTTGCAGTAATATTTTCAAAGATATATAGGATTTTACAAGAAAT	2325
QY	125	AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu	144
DB	2326	GATGATGCCAAACAAACGACAGAAATAGTAACTTGTTTCACTATTGTTCTAGTCCTT	2385
QY	145	GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu	164
DB	2386	GGAAATGTTCTTTTATTACATTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGAG	2445
QY	165	IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla	184
DB	2446	ATCCTCACCAAGCGCTCCGATACATACATGTTTCCGATCCATCCATCAGACAGGATGAGC	2505
QY	185	TyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle	204
DB	2506	TGGTTTGATAGCTTAAACACCACTGGAGCATTTGACTTACCAGCTCGCCAATGATGT	2565
QY	205	AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn	224
DB	2566	GCTCAAGTTAAGGGGCTATAGTTCCAGGCTGTGTAATTAATCCAGAAATATAGCAAT	2625
QY	225	MetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuLeuLeu	244
DB	2626	CTTGGCAGAGGAATAATTATATCTTAATCTATGTTGGCAACTGACACTGTTACTCTTA	2685
QY	245	SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe	264
DB	2686	GCAATTTGACCCATCATTCGAATAGCAGGAGTTGTGAAATGAAATGTTGTCTGGACAA	2745
QY	265	AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu	284
DB	2746	GCACGTGAAGATGAAGAAGAACTAGAGGTGCTGGAGAGATCGCTACTGAACAAATAGA	2805
QY	285	AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu	304
DB	2806	AACTTCCGAACCTGTTGTTTCTTGTGACTCAGGAGCAGAGTTTGAACATATGATGATCAG	2865
QY	305	MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr	324
DB	2866	AGTTTGCAGGTACCATACAGAAACCTTTTGAGGAAAGCACACATCTTTGGAATCAGTTT	2925
QY	325	AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr	344
DB	2926	TCCTTCAGCAGGCAATGATGATTTTTCCTATGCTGGATGTTTCGGTTTGGAGCCTAC	2985
QY	345	LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla	364
DB	2986	TTGGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTGTC	3045
QY	365	TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys	384



Qy	425	PhetYrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg	444
Db	3235	AACTATCCACCCGACTGCACATCCCGAGTTCACGGGCTGAGCCTCGAAGAAAG	3294
Qy	445	GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu	464
Db	3295	GGCCAGACGCTGCCTGTGTGGCAGCAGTGGCTGTGGAGAGACGCGTGTCCAGCTC	3354
Qy	465	LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys	484
Db	3355	CTGAGCGGTCTTATACCCCTTGGCGGGGAAGTGCCTGTTCAGCGCAAGAATAAAG	3414
Qy	485	GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu	504
Db	3415	CAACTGATGTTTCAGTGGCTCCGAGCACACCTGGGCGCATCGTGTCCAGAGACCATCCTG	3474
Qy	505	PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu	524
Db	3475	TTTGACTGCAGCATTAGTGAGAACATTGCCTATGAGACAACAGCGCGGTGTGTCCACAG	3534
Qy	525	AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro	544
Db	3535	GAAGAGATCGTGTGAGGCGACCCAGAGAGGCCAATATACAGCCTTCATCGAGTCACTGGCT	3594
Qy	545	GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln	564
Db	3595	AATATATATAGCACACGAGTAGGAGACAAGGAACCTAGCTCTCTGTGTGGCCAGAACAA	3654
Qy	565	ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla	584
Db	3655	CGCATTTGCCATAGCTCGTCCCTTGTGTACAGACCTCATATTATTTGCTTTTGGATGAAGCC	3714
Qy	585	ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg	604
Db	3715	ACATCAGCTCTCGATACAGAAAGTGAAGGTTGTCCAGAGAGCCTCGACAAAGCCAGA	3774
Qy	605	ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu	624
Db	3775	GAAGGCCGTACCTTGCATTTGTGATTGCTACCCGCCCTGTCCACCATCCAGAAATGCAGACTTA	3834
Qy	625	IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg	644
Db	3835	ATAGTGTGTGTTTCAGAAATGGCAGAGTCAGAGAGCAGGCACACATCAGCAGCTGCTGC	3894
Qy	645	AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	657
Db	3895	CAGAAAGGCATCTATTTTCAATGGTCACTGTCCAGGCT	3933

AAQ70907  
ID AAQ70907 standard; DNA: 2726 BP.

AC AAQ70907;

XX Multidrug-resistance gene MDR-1.  
DE

KW Multidrug-resistance; MDR-1 gene; ds.

OS Synthetic.

PN  
YY  
WO9417206-A.

XX  
SECRET

XX

PR 16-DEC-1993; 93US-0168621.

FA (UNCO-) UNCOX INC.

**Bhatnagar SK, George AL;**

**WPI; 1994-264118/32.**

Enzymatic amplification of target nucleic acid sequences in a mixt. - to detect a mutation or allele in the target, e.g. that causes a genetic disease

Disclosure; Page 30; 50pp; English.

This gene is used as an example of a target DNA in a method for the amplification of nucleic acid sequences in a mixture using various DNA probes/primers. The method may be used to detect a mutation or CC allele in the target, e.g. that causes a genetic disease.

Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

Alignment Scores:

Pred. No.:	7.82e-182	Length:	2726
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	15	Gaps:	4

US-09-873-409-1 (1-659) x AAQ070907 (1-2726)

Yq	1	MetLeuAlaGluLysGlyAHisAlaGluLeuMetalAlaLysArgGlyLeuTyrSer	20
Dd	317	GTCAITTTGGGAANGAAATCATGATAACTCATGAAGAGAAGCGCTTACITCAA	376
Qy	21	LeuValMetSerGin-----AspileLylysAlaSpGluGlnMet	34
Dd	377	CTTGTCACAATGCAGACAGCAGGAATGAAGTTGAATTAGAAAATGCAGCTGATG	436
Qy	35	GluSerMetThrTyrSerThrGluArgLysThrAasSerLeuProLeuHissSerVal---	53
Dd	437	AAAAAGTGAATTGATGCCCTTGGAAATGCTTCCAATGATTTCAAGATCCAGCTTAATAAGA	496
Qy	54	-----LysSerIleLys--SerAspPheIleAspLysAlaGluGluSer	67
Dd	497	AAAAGATCAACTGTAGGAGTGTCCTGGTGAATCAAGCCCCAACAGACAAGCTTAGTACC	556
Qy	68	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAan	87
Dd	557	AAAGAGCGCTCGGATGAAGATATACCTCCAGTTCCTTTTTGGAGGATTATGAGCTAAT	616
Qy	88	LysProGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrVal	107
Dd	617	TAACTGAATGGCCTTATTTTGTTGGTGATATTTGTGCCATTATAATGAGGCGCTG	676
Qy	108	HisProValPheSerIlePheAlaLysIleileThrMetPheGlyAan---AanAsp	126
Dd	677	C AACGACGATTGCAATAATATTTTCAAGAGATTATAGGGTTTTTACAGAAATTGAAT	736
Qy	127	LysThrThrLeuLysHisAspAlaGluIleTy-SerMetIlePheValIleLeuGlyVal	146
Dd	737	CCTGAACAAAACACAGACAGATAGTAACCTGTTTCACTATTGTTCTAGCCCTTGAAT	796
Qy	147	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaglyGluIleLeu	166
Dd	797	ATTTCTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC	856
Qy	167	ThrMetArgLeuArgHisLeualaPheIysAlaMetLeuTyrGlnAspIleAlaTrpPhe	186
Dd	857	ACCAAGCGCTCCCATACATGGTTTTCOGATCCATGCTCACAGAGGATGTAGTGGTTT	916
Qy	187	AspGluLysGluAanSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	206
Dd	917	GATGACCCCTAAAACACACCATCTGGAGCAATGATACCAAGGCTCGCCAAATGATGCTCTCAA	976
Qy	207	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	226

```
Db 2057 GCCATAGCTGCGCCCTTGTAGACAGCCTCATATTTTGTGATGAAGCCAGCTCA 2116
Qy 587 AlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 2117 GCTCTGGATACAGAAAGTGAAGAGTGTGTCGAAGAGCCTGGACAAAGCCAGAGAAGC 2176
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaSpLeuIleVal 526
Db 2177 CGCACCTGCATTGTGATTGTCTCACCGCTGTCCACCATCCAGAAATCCAGACTTAATAGTG 2236
Qy 627 ValLeuHisAsnGlyLysIleLysGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 2237 GTGTTTCAGAAATGGCAGAGTCAGAGGATGCGCACCATCAGCAGCTGTCTGGCAGAGAA 2296
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 2297 GGCATCTATTTTCAATGGTCAGTGTCCAGGCT 2329

RESULT 11
AAQ070916
ID AAQ070916 standard; DNA; 2726 BP.
XX
AC AAQ070916;
XX
DT 27-MAR-1995 (first entry)
XX
DE Multidrug-resistance gene MDR-1.
XX
KW Multidrug-resistance; MDR-1 gene; ds.
XX
OS Synthetic.
XX
PN WO9417210-A.
XX
PD 04-AUG-1994.
XX
PF 25-JAN-1994; 94WO-US00748.
XX
PR 27-JAN-1993; 93US-0010433.
PR 16-DEC-1993; 93US-0168621.
XX
PA (ONCO-) ONCOR INC.
XX
PI Bhatnagar SK, George AL;
XX
DR WPI; 1994-264122/32.
XX
PT Enzymatic amplification of target nucleic acid sequences to form
PT exact or modified copies - has increased fidelity and can identify
PT point mutations or allele(s)
XX
PS Disclosure; Page 30; 69pp; English.
XX
CC This gene is used as an example of a target DNA in a method for the
CC amplification of nucleic acid sequences in a mixture using various
CC DNA probes/primers. The method may be used to detect a mutation or
CC allele in the target, e.g. that that causes a genetic disease.
XX
SQ Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

Alignment Scores:
Pred. No.: 7,82e-182 Length: 2726
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 15 Gaps: 4

US-09-873-409-1 (1-659) x AAQ070916 (1-2726)
Qy 1 MetLeuAlaGlyLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 317 GTCAITGTGGAGAAAGGAATCAATGATGAACATCATGAAGAGAAAGGCATTACTTCAAA 376
```

```
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 377 CTTGTCAACATGCAGACAGCAGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 436
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 437 AAAAGTGAATGTAGTCCCTTGGAAATGCTTCAAAATGATTCAGATCCAGTCTAATAAGA 496
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCACAGAAAGCTTAGTACC 556
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 557 AAAAGAGGCTCTGGATCAAAAGTATACCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAAT 616
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 617 TTAAGTGAATGGCCTTATTTGTTGGTGTATTTGTCATTTATTAATGAGGCCTG 676
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 677 CAACCCAGCATTTGCAATAATATTTTCAAGATATATAGGGTTTTCACAAGAATTGATGAT 736
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 737 CCTGAAACAAAACGACAGATAGTAACCTTGTTCACCTATTTCTAGCCCTTGGAAAT 796
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 797 ATTTCTTTATTTACATTTTCTTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 856
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 857 ACCAAGCGGCTCCGATACATGCTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 916
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 917 GATGACCCCTAAAACACCACTGGAGCATTTGACTACAGGCTCGCCAATGATGTGCTCAA 976
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 977 GTTAAGGGGCTATAGGTTCCAGGCTTCTGTAATTTACCAAGATATAGCAATCTTGGG 1036
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
Db 1037 ACAGGAATAATTATATCTTCTATGTTGGCAACTAACACTGTACTCTTAGCAATT 1096
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 1097 GTACCCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGTCTGCAAGCACTG 1156
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 1157 AAAGATAAGAAAGAACTAGAGGTGCTGGAAAGATCGCTACTGAAGCAATAGAAACTTC 1216
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 1217 CGAACCGTTGTTTCTTGTGACTCAGGAGCAGAGATTGTAACATATGATGCTCAGAGTTTG 1276
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 1277 CAGGTACCATACAGAAACTCTTTTGAAGAAAGCACACATCTTTTGGAAATTACATTTTCTTC 1336
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db 1337 ACCCAGGCAATGATGATTTTTCCTATGCTGGAATGTTTCCGGTTTGGAGCCCTACTTGGG 1396
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTGTAGTATTTTTCAGCTGTTGTCTTGGT 1456
```



QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
 Db 1457 GCCATGCGCGTGGGCAAGTCAGTTCATTTGCTCTGACATATGCCAAAGCCAAATATCA 1516  
 QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysProAsnIleAspSerArgSerGlnGlu 406  
 Db 1517 GCAGCCACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1576  
 QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426  
 Db 1577 GGCTAATGCGCAACACATGGAAGGAATGTCATTTGGTGAAGTTGTATTCAACTAT 1636  
 QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
 Db 1637 CCCACCCGCGGACATCCAGTCTTCAGGACTGAGCTGGAGGTGAAGAGGGCCAG 1696  
 QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
 Db 1697 ACGTGGCTCTGTGGGCGAGCGTGGCTGTGGAGAGCACAGTGGTCCAGCTCTGTGAG 1756  
 QY 467 ArgLeuTyrPheProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
 Db 1757 CGGTTCACGACCCCTTGGCAGGAGAGTGTCTGTGATGCAAGAAATTAAGCGACTG 1816  
 QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
 Db 1817 AATGTTCACTGGCTCCGAGCACCTGGGCATCGTGTCCAGGAGCCCATCTGTTTGAC 1876  
 QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526  
 Db 1877 TGCAGCATTCGTGAGAACATTCCTATGGAGACACAGCCGGGTGTCTCACAGGAAGAG 1936  
 QY 527 IleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
 Db 1937 ATCGTGGGCGAGCAAGAGGCGCAACATACATGCTTCATCGAGTCACTGCCTAATAAA 1996  
 QY 547 TyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
 Db 1997 TATAGCACTAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGCAACACGCATT 2056  
 QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586  
 Db 2057 GCCATAGCTCGTCCCTTGTAGACAGCTCATATTTTGTCTTTTGGATGAAGCCACGTCA 2116  
 QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
 Db 2117 GCTCTGGATACAGAAAGTAAAGGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGGC 2176  
 QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
 Db 2177 GGCACCTGCATTTGTGATGCTCACCGCTGTCTCCACATCCAGATCCAGACTTATATAGT 2236  
 QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646  
 Db 2237 GTGTTTCAGATGGCAGAGTCAAGGAGCATGCGCATCGCATCGAGCTGTCTGTCACAGAA 2296  
 QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657  
 Db 2297 GGCATCTATTTTCAATGCTGAGTGTCCAGGCT 2329  
 RESULT 12  
 AAT43322  
 ID AAT43322 standard; DNA; 2726 BP.  
 XX  
 AC AAT43322;  
 XX  
 DT 01-SEP-1997 (first entry)  
 XX  
 DE Multidrug resistance gene-1.  
 XX  
 KW PCR; primer; amplify; polymerase chain reaction; ligase chain reaction;  
 KW LCR; human; multidrug resistance gene; MDR-1; ds.  
 XX

OS Synthetic.  
 XX WO9639537-A1.  
 XX  
 PD 12-DEC-1996.  
 XX  
 PF 04-JUN-1996; 96WO-US08841.  
 XX  
 PR 05-JUN-1995; 95US-0461823.  
 XX  
 XX (ONCO-) ONCOR INC.  
 XX  
 PI Bhatnagar SK, George AL, Nazarenko I;  
 XX  
 DR WPI; 1997-043158/04.  
 XX  
 PT Amplification method avoiding strand displacement by polymerase -  
 PT used in the detection of mutation(s) and allele(s) associated with  
 PT genetic disease and cancer  
 XX  
 PS Example 1; Page 50-51; 92pp; English.  
 XX  
 CC This sequence represents the multidrug resistance gene (MDR-1) amplified  
 CC by the primers shown in AAT43320 and AAT43321. This sequence can be used  
 CC as a target in the method of the invention, for enzymatically amplifying  
 CC a target nucleic acid (TNA) sequence contained in a nucleic acid or  
 CC mixture of nucleic acids while avoiding strand displacement by  
 CC polymerase. The method comprises using three primers, one complementary  
 CC to a first segment of the TNA, a second complementary to a second segment  
 CC of the TNA, which is adjacent to the first primer, and a third which is  
 CC similar to the first segment of the TNA, and is complementary to a  
 CC portion of the first primer. The first two primers are hybridised to the  
 CC TNA, and a fused amplification product is created from the TNA using  
 CC these two primers. The fused product is dissociated, and hybridised to  
 CC the third primer, which is then extended. The extended modified  
 CC amplification product is hybridised to the first two primers, and these  
 CC two primers are ligated. Each of the primers may be labelled using a  
 CC different label so that the method can be used to detect the presence of  
 CC a mutation or allele by detecting whether the labelled primer is  
 CC contained within the fused amplification product or the extended  
 CC amplification product. The method of the invention combines certain  
 CC aspects of ligase chain reaction (LCR) and polymerase chain reaction  
 CC (PCR), but is improved compared to LCR due to the reduced number of  
 CC primers needed, and the fact that the entire TNA sequence does not need  
 CC to be known.  
 XX  
 SQ Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7.82e-182 Length: 2726  
 Score: 1872.00 Matches: 356  
 Percent Similarity: 74.22% Conservative: 142  
 Best Local Similarity: 53.06% Mismatches: 159  
 Query Match: 56.33% Indels: 14  
 DB: 18 Gaps: 4  
 US-09-873-409-1 (1-659) x AAT43322 (1-2726)  
 QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
 Db 317 GTCATTGTGGAGAAAGAAATCATGATCACTCAAGAAAGAGAAAGGATTTACTTCAA 376  
 QY 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34  
 Db 377 CTTGTCAATGCAGACAGCAGGAAATGAAGTTGAATTAAGAAATGCAGCTGATCAATCC 436  
 QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
 Db 437 AAAAGTGAATTTGATGCTCTGGAATGCTTCAATGATTCATCAAGATCCAGTCTAATAAGA 496  
 QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67  
 Db 497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAGCCCAAGCAAGAAAGCTTAGTACC 556



XX Kreutzer R, Limmer S, Rost S, Hadwiger P;  
 XX WPI; 2002-270454/32.  
 XX  
 PT Inhibiting gene expression in cells, useful for e.g. treating tumors,  
 PT by introducing double-stranded complementary oligonucleotides having unpaired  
 PT terminal bases -  
 XX  
 PS Claim 13; Page 28-30; 104pp; German.  
 XX  
 CC The invention relates to a method for inhibiting expression of a target  
 CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
 CC oligoribonucleotide that has a double-stranded structure consisting of at  
 CC most 49 sequential nucleotide pairs, with at least part of one strand  
 CC complementary with the target gene and has at least one end a  
 CC single-stranded segment of 1-4 nt. The method provides  
 CC oligoribonucleotides for antisense inhibition of gene expression useful  
 CC e.g. for treating tumors but the oligoribonucleotides may also be  
 CC directed against genes present in pathogens (e.g. Plasmodium or  
 CC viruses/viroids, pathogenic on humans, animals or plants) or against  
 CC cytokine, Id, developmental or prion genes. The method provides more  
 CC effective inhibition of gene expression than use of known  
 CC oligonucleotides, probably because the unpaired overhang increases  
 CC stability and thus intracellular concentration.  
 XX  
 SQ Sequence 3840 BP; 1130 A; 738 C; 957 G; 1015 T; 0 other;  
 XX

Alignment Scores:  
 Pred. No.: 1,31e-181 Length: 3840  
 Score: 1872.00 Matches: 356  
 Percent Similarity: 74.22% Conservative: 142  
 Best Local Similarity: 53.06% Mismatches: 159  
 Query Match: 56.33% Indels: 14  
 DB: 24 Gaps: 4

US-09-873-409-1 (1-659) x ABL91687 (1-3840)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgLysLeuTyrTrpSer 20  
 Db 1810 GTCAATGTGGAGAAAGAAATCATGTGACTCATGAAAGAAAGGCAATTTACTTCAA 1869  
 Qy 21 LeuValMetSerGln-----AspLysLysAlaAspGluMet 34  
 Db 1870 CTTGTCACAATGCACAGCAGCAAGAAATGAATGAAATAGAAATGCAGCTCATGAATCC 1929  
 Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
 Db 1930 AAAAGTGAATGTAGTCCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 1989  
 Qy 54 -----LysSerileLys---SerAspPheileAspLysAlaGluSer 67  
 Db 1990 AAAAGATCAACTCGTAGGAGTGTCCGTGATCAAGCCCAAGCAAGAGCTTAGTACC 2049  
 Qy 68 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLeuLysLysLeuAsn 87  
 Db 2050 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTCTCTTTGGAGGATATGAAGCTAAAT 2109  
 Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
 Db 2110 TPAACGTGAATGGCCCTTATTTTGTGTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 2169  
 Qy 108 HisProValPheSerilePheAlaLysLysLysLysLysLysLysLysLysLysLysLys 126  
 Db 2170 CAACGAGCATTGGCAATATATTTTCAAGATATATAGGGGTTTTCACAGAAATGATGAT 2229  
 Qy 127 LysThrThrLeuLysHisAspAlaGluLysLysLysLysLysLysLysLysLysLysLys 146  
 Db 2230 CCTGAAACAAACGACAGATAGTAACTTGTGTTTCTTACATTTGTTTCTAGCCCTTGAAT 2289  
 Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuLeu 166  
 Db 2290 ATTTCTTTTATTTACATTTTCTTCTCAGGGTTCACATTTTGGCAAGCTGGAGAGATCCTC 2349

Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186  
 Db 2350 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACGATGTGAGTTGGTTT 2409  
 Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
 Db 2410 GATGACCCCTAAACACACCACCTGGAGCATGTACACAGGCTCGCCCAATGATGCTCTCAA 2469  
 Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
 Db 2470 GTTAAAGGGGCTATAGGTTCCAGGCTTCTGTATATACCAGAAATATAGCAATCTTGGG 2529  
 Qy 227 LeuSerValIleLysSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerile 246  
 Db 2530 ACAGGAATAATATATATCTTCTATGTTGGCAACTAACACACTGTTACTCTTAGCAATT 2589  
 Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
 Db 2590 GTACCCATCATTTGCAATAGCAGGAGTTGTTGAATGAAATGTTGCTGGGACACGACTG 2649  
 Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
 Db 2650 AAAGATAAGAAAGAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTC 2709  
 Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
 Db 2710 CGAACCGTTGTTCTTCTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTG 2769  
 Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
 Db 2770 CAGGTACCATACAGAAACTCTTTGAGAAAGCAGACATCTTTGGAAATTTACATTTTCCCTC 2829  
 Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIle 346  
 Db 2830 ACCGAGCAATGATGATTTTCTATGCTGATGTTTCCGTTTGGAGCCCTACTTGTGTG 2889  
 Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
 Db 2890 GCACATAAATCATGAGCTTTGAGGAGTCTCTGTTAGTATTTTTCAGCTGTTGCTCTTGGT 2949  
 Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
 Db 2950 GCCATGGCCGTGGGCAAGTCAGTTTCATTTGCTCTGATGTCATGCCAAAGCCCAAAATATCA 3009  
 Qy 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
 Db 3010 GGAGCCACATCATCATGATCATTTGAAAGAAACCCCTTTGATTGACAGCTACAGCAGGAA 3069  
 Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426  
 Db 3070 GGCTTAATGTCGAAACACATTTGAGAAATGTTCACATTTGTTGTAAGTTGTTATTCAACTAT 3129  
 Qy 427 ProCysArgProAspValPheIleLeuArgLysLeuSerLeuSerileGluArgLysLys 446  
 Db 3130 CCCACCCGACGGACATCCCATGCTTCCAGGACCTGAGCCCTGGAGGTGAAGAGGCGCCAG 3189  
 Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466  
 Db 3190 ACGTGGCTCTGTTGGGAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGTCTCTGGAG 3249  
 Qy 467 ArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeu 486  
 Db 3250 CGGTCTACAGACCCCTTGGCAGGAAAGTGTGCTTGTATGTCGCAAGAAATTAAGGACCTG 3309  
 Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
 Db 3310 AATGTTCACTGGCTCCGAGCACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTTAC 3369  
 Qy 507 CysSerileAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526  
 Db 3370 TGCAGCATTTGCTGAGAAATGTCCTATGAGACAAACAGCCGGGTGTTGTACAGGAAGAG 3429



```
Db 2293 ATTTCTTTTATACATTTTCTCTCAAGGTTTCACATTTGGCAAGCTGGAGAGATCTCTC 2352
Qy 167 ThrMetArgLeuAtrGHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 2353 ACCAAGCGGCTCCGATACATGGTTTTCGATCCATGCTCAGACAGGATGTGATGGTTT 2412
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2413 GATGACCCCTAAAAACACCACTGGAGCATTTGACTACAGGCTCGCCAAATGATGCTGCTCAA 2472
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2473 GTTAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTTACCCAGAAATATAGCAAAATCTTGGG 2532
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 246
Db 2533 ACAGGAATAATATATATCTTCATCATGTTGGCAACTAACACTGTTACTCTTAGCAAT 2592
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 2593 GTACCACATTCGTAATAGCAGGATGTTGTAATGAAATGTAATGTTGCTGGACAAGCACTG 2652
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2653 AAAGATAAGAAAGAACTAGAGGTGCTGGGAAGATCGCTACTGAAAGCAATAGAAAACCTTC 2712
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 2713 CGAACCGTGTTCCTTGACTCAGGAGCAAGAGTTTGAACATATATGATGCTCGAGATTG 2772
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 2773 CAGGTACCATACAGAAACTCTTTGAGGAAGCACACATCTTTTGGAAATTACATTTTCCITTC 2832
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuLe 346
Db 2833 ACCCAGGCAATGATGATTTTCTCATGTGATGTTCCGGTTTGGAGCCCTACTTGGTG 2892
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 2893 GCACATAAACTACAGATCTTCAGAGATGTCGTGATGATTTTCAGCTGTTGCTCTTGTG 2952
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 2953 GCCATGGCGCTGGGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAAGCCAAATATCA 3012
Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3013 GCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3072
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426
Db 3073 GGCCCTAAATGCCCAACACATTTGAAGGAATGTGCACATTTGGTGAAGTTGATTTCAACTAT 3132
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3133 CCCACCCGACCCGACATCCAGTGTTCAGGAGCTGAGCTGGAGGTGAAGAGGGCCAG 3192
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 3193 ACGTGGCTCTGTGGCCAGCAGTGGCTGTGGGAAGAGCACAGTGTGCTCCAGCTCTCTGGAG 3252
Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3253 CGGTTCTACGACCCCTTCGACGGGAAAGTGTGCTGTGATGGCAAAATAAAGCACTG 3312
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3313 AATGTTCACTGGCTCCGACACACCTGGGCATCGTGTCCAGGAGGCCATCTCTGTTGAC 3372
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
```

```
Db 3373 TGCAGCATTTGCTGAGAACATTTGCCCTATGGAGACAACAGCCGGTGGTGTCTACAGNAGAG 3432
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3433 ATCGTGGGGCAGCAAGAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTAATAAA 3492
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 3493 TATAGCACTAAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACACGCAAT 3552
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3553 GCCATAGCTGTGGCTTTGTAGCAGCTCATATTTTGTCTTTGATGAAGCCACGTCAC 3612
Qy 587 AlaLeuAspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3613 GCTCTGGATACAGAAAGTGAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGGC 3672
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 3673 CGCACCTCATTTGATTTGCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTG 3732
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646
Db 3733 GTGTTTTCAGAAATGGCAGAGTCAAGGAGCATGCGACCATCAGCATCAGCTGCTGCGCACAGAAA 3792
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 3793 GGCATCTATTTTCAATGTCAGTGTCCAGGCT 3825
RESULT 15
AAZ49333
ID AAZ49333 standard; cDNA; 3860 BP.
XX
AC AAZ49333;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX
KW Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
XX cytokine; mutant; ds.
OS Synthetic.
OS Homo sapiens.
XX
Key Location/Qualifiers
CDS 1..3843
FT /tag= a
FT /product= "Human G185V mutant MDR-1 protein"
XX
XX WO9961589-A2.
XX
XX 02-DEC-1999.
XX
XX 27-MAY-1999; 99WO-US11825.
XX
XX 28-MAY-1998; 98US-0086988.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Sorrentino B, Bunting K;
XX
XX WPI; 2000-072615/06.
XX
XX P-PSDB; AAY58187.
XX
XX Ex vivo expansion of hematopoietic stem cells transduced with a
XX sequence encoding human multidrug resistance-1, used for bone marrow
XX transplantation -
PT
```

XX Example 1; Page 79-82; 113pp; English.

XX This sequence represents cDNA encoding human G185V mutant multidrug  
 CC resistance protein MDR-1, where the Gly residue at position 185  
 CC of the wild-type protein (P455186) is replaced by Val. MDR-1 is a  
 CC transmembrane efflux pump, responsible for the export of drugs from  
 CC cells, particularly cancer cells. The wild-type MDR-1 shows increased  
 CC resistance to etoposide and decreased resistance to vinca alkaloids  
 CC compared with the G185V mutant. The invention relates to transducing  
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified haematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC haematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from haematopoietic stem cells, e.g., thalassemia,  
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in haematopoietic stem cells.  
 CC Haematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC haematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.

XX Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

#### Alignment Scores:

Pred. No.: 1,32e-181 Length: 3860  
 Score: 1872.00 Matches: 356  
 Percent Similarity: 74.22% Conservative: 142  
 Best Local Similarity: 53.06% Mismatches: 159  
 Query Match: 56.33% Indels: 14  
 DB: 21 Gaps: 4

US-09-873-409-1 (1-659) x AAZ49333 (1-3860)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTrpSer 20  
 DB 1813 GTCAATGTGAGAAAGAAATCATGATGATGAACTCATGAAAGAGAAAGCAATTTACTTCAA 1872  
 QY 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34  
 DB 1873 CTTGTCAACATGCAGACAGCAGGAAATGAAGTTAGAAATGCAAAATGCAGCTGATGAATCC 1932  
 QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
 DB 1933 AAAAGTGAATTCATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGA 1992  
 QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67  
 DB 1993 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAGCCCAAGCAAGCTTAGTACC 2052  
 QY 68 ThrGlnSerLysGluLysSerLeuProGluValSerLeuLysLysIleLeuLysLeuAsn 87  
 DB 2053 AAAAGAGCTCTGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTAAGAAGTAAAT 2112  
 QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuGlyThrVal 107  
 DB 2113 TTAACCTGAATGGCTTATTTGTGTGTGTATTTTGGCCATATAAATGGAGCCCTG 2172  
 QY 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126  
 DB 2173 CAACAGCATTTGCATATATATTTCAAGATATATAGGGGTTTACAGAAATGATGAT 2232  
 QY 127 LysThrLeuLysHisAspAlaGluLysThrSerMetIlePheValIleLeuGlyVal 146  
 DB 2233 CCTGAAACAAAACAGACAATAGTAACTTTGTTTCTACTATTGTTCTAGCCCTTGAAT 2292  
 QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeu 166

DB 2293 ATTTCTTTTATACATTTTCTTCAAGGTTTCAATTTGGCAAGCTCGAGAGATCCTC 2352  
 QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186  
 DB 2353 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2412  
 QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGln 206  
 DB 2413 GATGACCTCTAAAACACCACCTGGAGCATTTGACTACAGGCTGCCAATGATGCTCAA 2472  
 QY 207 IleGlnGlyValThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrMetGly 226  
 DB 2473 GTTAAAGGGGCTATAGTTCCAGGCTGCTGTAATTCACCAATATAGCAATCTTGG 2532  
 QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246  
 DB 2533 ACAGGAATAATTATATCTTCATCTATGGTGGCAACTAACACTGTTACTCTTAGCAAT 2592  
 QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
 DB 2593 GTACCCATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTGGACAAGCATG 2652  
 QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
 DB 2653 AAAGATAGAAAGAACTAGAGGCTGCTGGAAAGATCGCTACTGAAGCAATAGAAACTTC 2712  
 QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
 DB 2713 CGAACCGTTGTTTCTTGTGACTCAGGAGCAGAGTTTGAACATATGATGCTCAGAGTTG 2772  
 QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
 DB 2773 CAGGTACCATACAGAAACTCTTTGAGGAAAGCAGACATCTTTTGGAAATACATTTCTTC 2832  
 QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValatyrLeuIle 346  
 DB 2833 ACCCAGGCAATGATGATATTTTCTATGCTGGATGTTCCGGTTTGGAGCCTTACTTGGT 2892  
 QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
 DB 2893 GCACATAAATCATGAGCTTTGAGGATGTTCTGTGATGATTTTTCAGCTGTTCTTGTGT 2952  
 QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
 DB 2953 GCATGGCGCTGGGCAAGTCAGTTCAITTTGCTCTGACTATGCCAAAGCCAAATATCA 3012  
 QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
 DB 3013 CGAGCCCAATCATCATGATCATTTGAAAGAAACCCCTTTGATTCAGAGCTACAGCAGGAA 3072  
 QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426  
 DB 3073 GGCCTAATGCCGAACACATGGAAGAAATGTCACATTTGGTGAAGTTCTATTCAACTAT 3132  
 QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
 DB 3133 CCCACCCGCGGACATCCAGTGTCTCAGGAGCTGAGCTGGAGGTGAAGAGGCGCCAG 3192  
 QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466  
 DB 3193 ACGTGGCTCTGTTGGGAGCAGAGTGGCTGTGGGAAAGACACAGTGGTCCAGCTCTCTGG 3252  
 QY 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
 DB 3253 CGGTTCTACGACCCCTTGGCAGGGAAGTGTGCTTGTATGGCAAGAAATAAAGCGCATG 3312  
 QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlnProValLeuPheAsn 506  
 DB 3313 AATGTTAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTTAC 3372  
 QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAsnSerArgValValProLeuAspGlu 526  
 DB 3373 TGCACATTCGTGAGAACATTCGCTATGAGACAAACAGCCGGGTGGTGTACAGGAAGAG 3432

```
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3433 ATCGTAGGGCAGCAAGAGGCCAACATACATATGCTTCATCGAGTCACGCTAATAAA 3492

Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3493 TATAGCACTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3552

Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3553 GCCATAGCTCGTGCCCTTGTAGACAGCTCATATTTTGGTATGAAGCCACGTCA 3612

Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3613 GCTCTGGATACAGAAAGTGAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3672

Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3673 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCAATCCAGAAATGCAGACTTAATGTG 3732

Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3733 GTGTTTCAGATGGCAGAGTCAGAGGATGCACGCATCAGCAGCTGCTGGCACAGAAA 3792

Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3793 GGCATCTATTTTCAATGGTCAGTGTCAGGCT 3825
```

Search completed: March 30, 2003, 03:05:40  
Job time : 735.062 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 4532.72 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3323

Sequence: 1 MLEAKGAHAEALMAKRGGLYS.....QELLNRDIYFKLVNAQSVQ 659

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPFO\_spool/US09873409/runat\_27032003\_115420\_19240/app\_query.fasta\_1.7544  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 30544 @runat\_27032003\_115420\_19240 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_Other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1008	30.3	943	9	AL520322	AL520322 AL520322
2	834	25.1	998	14	BM904842	BM904842 AGENCOURT
3	780	23.5	2676	11	AK014319	Mus muscu
4	770.5	23.2	1019	12	BG248052	BG248052 602359987
5	753	22.7	726	12	BG293345	BG293345 602350738
6	751	22.6	944	12	BF796582	BF796582 602258463
7	743.5	22.4	886	17	BH139685	BH139685 ENTNA47TR
8	740.5	22.3	871	17	AZ682350	AZ682350 ENTKB16TF
9	734.5	22.1	932	17	AZ670821	AZ670821 ENTJN69TF
10	728.5	21.9	947	17	AZ683753	AZ683753 ENTJL96TF
11	726.5	21.9	886	17	AZ540627	AZ540627 ENTEQ18TF
12	723.5	21.8	880	17	AZ687805	AZ687805 ENTLS52TF
13	721.5	21.7	939	14	BQ720763	BQ720763 AGENCOURT
14	720.5	21.7	897	17	AZ541090	AZ541090 ENTDS67TR
15	711.5	21.4	913	17	BH155700	BH155700 ENTROS47TR
16	709.5	21.4	834	17	AZ548312	AZ548312 ENTFG07TR
17	708.5	21.3	853	17	AZ679807	AZ679807 ENTH167TR
18	705.5	21.2	823	17	AZ532602	AZ532602 ENTCS50TR
19	697.5	21.0	861	14	BQ717101	BQ717101 AGENCOURT
20	691	20.8	780	12	BG587938	BG587938 EST489713
21	690.5	20.8	899	17	BH154857	BH154857 ENTRO47TF
22	689.5	20.7	891	17	AZ682250	AZ682250 ENTKT68TR
23	688	20.7	1123	11	AY108485	AY108485 Zea mays
24	683.5	20.6	1341	11	AY108285	AY108285 Zea mays
25	683	20.6	600	13	BM486593	BM486593 pgm2n.pk0
26	682.5	20.5	926	17	AZ690701	AZ690701 ENTIV68TR
27	678	20.4	926	14	BQ123477	BQ123477 EST609053
28	677.5	20.4	795	12	BF313560	BF313560 601900192
29	669	20.1	547	13	BQ307011	BQ307011 BQ307011
30	667	20.1	937	17	BH150760	BH150760 ENTQ42TR
31	663	20.0	636	13	BJ349604	BJ349604 BJ349604
32	662	19.9	636	10	AV962688	AV962688 AV962688
33	661	19.9	747	13	BJ445774	BJ445774 BJ445774
34	659.5	19.8	628	12	BG080311	BG080311 H3052B06-
35	659	19.8	852	17	BH720383	BH720383 BOH2294TF
36	657	19.8	715	14	BQ869512	BQ869512 QGD609.Y
37	657	19.8	785	10	AV709991	AV709991 AV709991
38	655.5	19.7	787	12	BG584063	BG584063 EST485823
39	653	19.7	609	10	AV986144	AV986144 AV986144
40	651.5	19.6	921	17	AZ687628	AZ687628 ENTIR36TR
41	651	19.6	746	13	BJ380617	BJ380617 BJ380617
42	650	19.6	729	13	BJ353293	BJ353293 BJ353293
43	647	19.5	1813	11	U66688	U66688 Homo sapien
44	641	19.3	820	13	BJ356540	BJ356540 BJ356540
45	639.5	19.2	842	17	AZ671925	AZ671925 ENTMN93TR

# ALIGNMENTS

RESULT 1  
AL520322  
LOCUS AL520322 LTI NFL004 NBC2 Homo sapiens CDNA clone CS0DB06YC15 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL520322 GI:12783815  
VERSION AL520322.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 943)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization

JOURNAL  
COMMENT

Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

FEATURES  
source

Location/Qualifiers  
1..943  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DB006YC15"  
/clone\_lib="LRI\_NFL004\_NBC2"  
/sex="male"  
/tissue type="neuroblastoma cells"  
/lab host="DH10B"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [fliang@life.com](mailto:fliang@life.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT 253 a 225 c 253 g 211 t 1 others

## ORIGIN

Alignment Scores:  
Pred. No.: 3.17e-110 Length: 943  
Score: 1008.00 Matches: 194  
Percent Similarity: 80.40% Conservative: 48  
Best Local Similarity: 64.45% Mismatches: 58  
Query Match: 30.33% Indels: 1  
DB: 9 Gaps: 0

US-09-873-409-1 (1-659) x AL520322 (1-943)

Qy 358 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLeuValLeuAla 377  
Db 5 TTATTATTTTTCAGCTGTTGCTTTGGTGCATGGCGGCGCAAGTCAGTTTCATTGCT 64  
Qy 378 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLys 397  
Db 65 CTTGACTATGCCAAGCCAAAATATCAGCAGCCCAATCATCATGATCATTTGAAAAACC 124  
Qy 398 ProAsnIleAspSerArgSerGlnGlnGlyLysLysProAspThrCysGluCysLeu 417  
Db 125 CCTTTGATTGACAGCTACAGCAGCGAAGCCCTTAATGCCGAACACATTTGGAAGGAATGTC 184  
Qy 418 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 437  
Db 185 ACATTGGTGAAGTTGATTCAACTATCCACCACCGACATCCAGTGTTCAGGGA 244  
Qy 438 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 457  
Db 245 CTGAGCTGGAGGTGAAGAGGCCAGACGCTGCTGTGTGGCAGCAGTGGCTGTGGG 304  
Qy 458 LysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlnValLeu 477  
Db 305 AAGAGCAGACGTGCTCAGCTCTGAGGCGGTTCACGACCCCTTGGCAGGGAAGTGTG 364  
Qy 478 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrrLeuArgSerGlnIleAlaIle 497  
Db 365 CTTGATGCCAAGAATAAAGCGACTGAATGTTCACTGCTCCGAGCACACCTGGGCATC 424  
Qy 498 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 517  
Db 425 GTGTCCAGGAGCCCATCTGTTTGTACTGTCAGCAATTCGTGAGCAATTCCTATGAGAC 484  
Qy 518 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHis 537  
Db 485 AACAGCCCGGGTGTGTCAGGAAGAGATCTGTGAGGGCAGCAAGAGGCGCAACATACAT 544

Qy 538 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 557  
Db 545 GCCTTCATCGAGTCAGTCCCTTAATAAATATAGCACTAAAGTAGGAGACAAGGAACCTCAG 604  
Qy 558 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 577  
Db 605 CTCTCTGTGTGGCCAGAAACAACGCAATTCGCATAGCTCGTGCCTTGTGTAGACAGCCTCAT 664  
Qy 578 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 597  
Db 665 ATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAA 724  
Qy 598 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 617  
Db 725 GAAGCCCTGGACAAGCCAGAGAAGCGGCACCTGCTGTTGTTGCTCACCCTGTGCC 784  
Qy 618 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGln 637  
Db 785 ACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGAATGCGCAGAGTCAAGGAGCATGG 844  
Qy 637 YThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 657  
Db 845 CAGCATCAGCAGCTGCTGGCAGCAAGAGGCATCTATTTTCAATGTCAGTGTCCAGGC 904  
Qy 657 r 657  
Db 905 t 905

## RESULT 2

BM904842

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

## FEATURES

source

1..998

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5557655"

/clone\_lib="NIH MGC 72"

/tissue type="melanotic melanoma"

/lab host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;

Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2 Kb. Library constructed by Life

Technologies."

BASE COUNT 271 a 223 c 248 g 256 t

ORIGIN

Alignment Scores:

Pred. No.: 3.26e-89 Length: 998

Score: 834.00 Matches: 164

[illegible]

LOCUS	AK014319	2676 bp	mRNA	linear	HTC 19-JAN-2000
DEFINITION	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:322401p09:ATP-binding cassette, sub-family B (MDR/TAP), member 8, full insert sequence.				
ACCESSION	AK014319.1 GI:12852089				
VERSION	HTC; CAP trapper.				
KEYWORDS	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library clone:322401p09.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1				
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	92729253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Harada,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Asiburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Tomita,M., Quackenbush,J., Schrim,L.M., Staabli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Balderelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Guetincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085560				
PUBMED	11217851				
REFERENCE	5 (bases 1 to 2676)				
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Balderelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Oono,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,				

RESULT 3  
AK014319



```
Db 1200 TTCTATGGGGCTCCCTGTGGCTGGACAGAGCTGAAAGGGGAGAGCCTCATGTCTCTTC 1259
Qy 358 lIeValPheThrAlaIleAlaTyrglyAlaMetAlaIleGlyLysThrLeuValLeuAla 377
Db 1260 CTGTGGCTTCCAGACAGTACAGAGGCTATGGCC-----AGCCTCTCTGTCTGT 1313
Qy 378 ProGluTyrsrLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 397
Db 1314 GGTcAGGTGGTACGTGGGCTGAGTGGGAGCCCGAGCTTCGGAATACATGCGCTGAGC 1373
Qy 398 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 417
Db 1374 CCTGTATCCATCCATGACCGGGGGCTACTGTCATCCCAACAAAGGACATTCGTGGTTCATC 1433
Qy 418 GluPheArgGluValSerPhePheTyrrProCysArgProAspValPheIleLeuArgGly 437
Db 1434 ACCTTCCAAATGTCACCTTCAGCTACCCCTGCAGACTTGGCTTCAATGTCTCAGGAC 1493
Qy 438 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 457
Db 1494 TTCACCTGAAGTCCCTCTGGCAAGATTGTGGCTCTTGTGGGCCAGTCTGGGGAGGA 1553
Qy 458 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrrAspProValGlnGlnValLeu 477
Db 1554 AAGACCAcAGTTCCTCTGCTGAACGCTTCTATGACCCCTGAAGCTGGCTCGGTGAGC 1613
Qy 478 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 496
Db 1614 TTGGATGGCGATGACCTCGCAACTCTCAACCCCTCTGGCTCCGGGGCCAGGTATAGGT 1673
Qy 497 lIeValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrrGly 516
Db 1674 TTCATCACCCAGGACCACTCTGTTTGCACACCACTCATGAGAAATATCCGATTGGG 1733
Qy 517 -----AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 534
Db 1734 AAGCTGGATGCTTCGAT-----GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781
Qy 535 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrrAsnThrGlnValGlyLeuLys 554
Db 1782 AATGCCCAcGAGTTTCATCAGCAGCTTCCCGATGGCTACAGCACTGTGGTGGTGAAGGG 1841
Qy 555 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGln 574
Db 1842 GGCACAACTTGTCTGTGGCCAGAGAGAGCGCTTAGCCATCGCAGTGCCTCATCAAG 1901
Qy 575 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLys 594
Db 1902 CAGCCCAcAGTCTGATCTGGACGAGGCCACcAGTGGCTAGATGCAGAATCCGAGAGG 1961
Qy 595 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 614
Db 1962 GTGGTACAGGAGGCGCTGGACCGGCCAGTGTCTGGCCGACCGCTGTGTGTCTATCCCCAC 2021
Qy 615 ArgLeuSerAlaIleGluAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 634
Db 2022 CGCTTAGTACTGTCTGGTCACCCCACTCCATCATTTGTATGGCCAAATGGCCAGTCTGT 2081
Qy 635 GluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrrPheLysLeuValAsn 654
Db 2082 GAGGTGGGACCCAGCAAGAACTCTTAAAGAGGGCGGCTCTATTTCAGAGCTTATCCGG 2141
Qy 655 AlaGlnSerVal 658
Db 2142 AGACAAACCCGT 2153
RESULT 4
BG248052 1019 bp mRNA linear EST 13-FEB-2001
LOCUS 602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
DEFINITION mRNA sequence.
ACCESSION BG248052
```

```
VERSION BG248052.1 GI:12757867
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1019)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10334 row: n column: 21
High quality sequence stop: 650.
FEATURES
source
1..1019
Location/Qualifiers
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488404"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 276 a 262 c 283 g 196 t
ORIGIN
Alignment Scores:
Pred. No.: 1,52e-81 Length: 1019
Score: 770.50 Matches: 154
Percent Similarity: 79.17% Conservative: 36
Best Local Similarity: 64.17% Mismatches: 47
Query Match: 23.19% Indels: 3
DB: 12 Gaps: 1
US-09-873-409-1 (1-659) x BG248052 (1-1019)
```

```
Qy 414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrrProCysArgProAspValPhe 433
Db 4 GAAGGAAATGTCAATTTAGTGGAGTCTGTTCAACTATCCCACTCCCAAGCAGCATCCCA 63
Qy 434 lIeLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
Db 64 GTGCTTCAGGGGCTGAGCTTTCAGGTGAAGAGGGCCAGACGCTGGCTTGTGGGAGC 123
Qy 454 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrrAspProValGln 473
Db 124 AGTGCTCGGGAGAGGACACAGTGTCTCAGCTGCTCGAGCGTCTTACGACCCCATGGT 183
Qy 474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
Db 184 GGATCAGTGTCTTAGATGGCAAGAAATAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243
Qy 494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
Db 244 CAGCTGGGCAATTGTCTCCCAAGAGGCCATTCTCTTTGACTGCAGCATCGCAGAGCAAT 303
Qy 514 AlaTyrrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 533
Db 304 GCCTACGGAGACAAcAGCGGGTCTGTCTTATAGGAGATTGTGAGGGCAGCCCAAGGAG 363
Qy 534 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrrAsnThrGlnValGlyLeu 553
```





High quality sequence start: 9  
 High quality sequence stop: 669.  
 Location/Qualifiers  
 source  
 1. 944  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4341710"  
 /clone\_lib="NIH MGC 85"  
 /tissue\_type="lymphoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.867 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH MGC Library."  
 BASE COUNT 265 a 241 c 237 g 201 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3e-79 Length: 944  
 Score: 751.00 Matches: 166  
 Percent Similarity: 74.62% Conservative: 31  
 Best Local Similarity: 62.88% Mismatches: 60  
 Query Match: 12.60% Indels: 9  
 DB: 12 Gaps: 2

US-09-873-409-1 (1-659) x BF796582 (1-944)

Qy 380 TyrSerLysAlaLysSerGlyAlaHisLeuPheAlaLeuLeuGluLysProAsn 399  
 Db 39 TATGCTAAAGCTAAGCTCTCTGCAGCCACCTATTATCATGCTTTTGAAGACACACCTCTG 98  
 Qy 400 IleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuPhe 419  
 Db 99 ATTGACACTACAGTGAAGGGGCTGAAGCTCTGTAATTTGAAGGAATATAACATTT 158  
 Qy 420 ArgGluValSerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSer 439  
 Db 159 AATGAAGTCGTGTTCAACTATATCCACCGAGCAACGTCGCCAGTCTTTCAGGGGTGAGC 218  
 Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459  
 Db 219 CTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCGAGCGTGGTGGGAAGAGC 278  
 Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyRhapProValGlnGlyGlnValLeuPheAsp 479  
 Db 279 ACGTGGTCCAGCTCTCGAGGGGTTCTACGACCCCTTGGCGGGGACAGTGTCTTCGAT 338  
 Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499  
 Db 339 GGTCAAGAGCAAGAAACTCAATCTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGCT 398  
 Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyRglyAspAsnSer 519  
 Db 399 CAGGAGCCTATCTTATTTGACTGCAGCATTTGCCGAGAATATTGCTTATGGAGACAAACAGC 458  
 Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539  
 Db 459 CGGGTTGTATCAGAGTGAATCTGTAGTCGAGCAAGCTGCCAACAATATCTCTTTC 518  
 Qy 540 IleGluGlyLeuProGluLysTyRAsnThrGlnValGlyLysGlyValAlaGlnLeuSer 559  
 Db 519 ATCGAGAGCTTACCCACAAATATGAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCA 578  
 Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLeu 579  
 Db 579 GGAGGTCAAAACAGAGGATTTGCTATTGCGGAGCCCTCATCAGACACACCTCAAAATCTC 638  
 Qy 579 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal-----G 597  
 Db 639 CCTGTTGATGAAGCTACATCAGCTCTGGATACCTGCAAGAGTGAACACGGTTGTCCTCAA 698  
 Qy 597 InHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 617

Db 699 GAAGCCCTGTACACAGACAGAGAGCGCCGCTGATGCTCACCCTGTC 758  
 Qy 617 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 637  
 Db 759 AAC--ATCCAGATGCAGACTT-ATAGTGGTGTTCCTCCCAACGGGAGAGTCAAG----- 808  
 Qy 637 lyThrHis 639  
 Db 809 --ACACAT 814  
 RESULT 7  
 BH139685 886 bp DNA linear GSS 07-AUG-2001  
 LOCUS ENTNA477R Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 DEFINITION genomic, DNA sequence.  
 ACCESSION BH139685.1 GI:15098746  
 VERSION BH139685  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 886)  
 Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.  
 AUTHORS Determination of clone end sequences from Entamoeba histolytica  
 TITLE HM1:IMSS sheared DNA library (2001)  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjlloftus@tigr.org  
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
 DNA library  
 Seq primer: M13-Reverse  
 Class: shotgun  
 High quality sequence start: 18  
 High quality sequence stop: 851.  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 886  
 /organism="Entamoeba histolytica"  
 /strain="HM1:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: pHOSt; Site 1: Bst I; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica  
 using a method described by Clark and Diamond (Clark,  
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
 method for isolate identification. Exp. Parasitol.  
 77:450.). The DNA was mechanically sheared to give a  
 tight size distribution (~2 kb). The v + i method used for  
 the library construction is described in detail in Smith,  
 H.O. and Venter, J.C. (Making small insert libraries for  
 whole genome shotgun sequencing projects. In Genome  
 Sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Borell, Oxford University Press, 1999)."  
 BASE COUNT 348 a 105 c 176 g 257 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.16e-78 Length: 886  
 Score: 743.50 Matches: 151  
 Percent Similarity: 73.43% Conservative: 48  
 Best Local Similarity: 55.72% Mismatches: 71  
 Query Match: 22.37% Indels: 2  
 DB: 17 Gaps: 1  
 US-09-873-409-1 (1-659) x BH139685 (1-886)  
 Qy 387 AlaAlaHisLeuPheAlaLeuLeuGluLysProAsnIleAspSerArgSerGlnGlu 406

```

Db 20 GCAAGATATCTATGATGTTATGATAGAAACCACT-ATTGATTTTATAGTGAGAA 78
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheThr 426
Db 79 GGTCAACATTTAATGATGTTAAAGGTGAATGAAITTAAGACATTTGTTTCAGATAT 138
Qy 427 ProCysArgProAspValPheLeuArgGlyLeuSerLeuSerLeuGluArgGlyLys 446
Db 139 CCAACAGACACACATCTGCTTGAAGGTATTTTCATTCAAAGTAGAACAGGAAAA 198
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466
Db 199 ACTGTTCATAGTAGAGCATCAGGATGTTGTAATCAACATCAGTTCAAGTTGATGAA 258
Qy 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAlaLysGluLeu 486
Db 259 AGATTTTATGATCCACACATGAGATGTTATTATAGATGACATAATCAAGATTG 318
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 319 AATATTCATTTCTAGAAAGTCAAAATGGATGGTAGGACAAAGACGATTTATTTGCT 378
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 525
Db 379 GAAAGTGTATGATATATATAGAGGAGTACCTAAAGGATTCAGTAAGTAATGAA 438
Qy 526 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 545
Db 439 CAAATTTATGCTGCTGCTAAATGGCAATGCACATGACTTTTATTCAGCAATGCCAGAA 498
Qy 546 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArg 565
Db 499 GGATATAACACATGTTAGGTGATGAGGTGCACAAATTTTCAGGAGCAAAAACAAAGA 558
Qy 566 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 585
Db 559 ATTGCTATTGCACGTGCTGATGATGAAATCCAAAGTGTATTCTCGATGAAGCTACA 618
Qy 586 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 605
Db 619 TCAGCACTTGATTCAGAAAGTGAAGATGTTACAAAGATGCACCTTCACAAAGCAGCAAA 678
Qy 606 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaSerLeuIle 625
Db 679 GGAAGAACCAACATTTGTAATTCACATAGATTATCAACTATTCAAAATGCAGATCAATA 738
Qy 626 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsn 645
Db 739 TGTGTTATTATGAGAGAGAAATTCAGAAAGAGGAAACATCATCAGAGTTATTAGATTG 798
Qy 646 ArgAspIleTyrPheLysLeuValAsnAlaGln 656
Db 799 AAAGGATTTTATTATACACTTGCTGCAACAA 831

```

```

RESULT 8
LOCUS AZ682350 871 bp DNA linear GSS 14-DEC-2000
DEFINITION ENTBK16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
GENOMIC, DNA sequence.
ACCESSION AZ682350
VERSION AZ682350.1 GI:11819496
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 871)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HMI:IMSS sheared DNA library
COMMENT Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics

```

The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjoftus@igr.org  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library  
 Seq primer: M13-Forward  
 Class: shotgun  
 High quality sequence start: 16  
 High quality sequence stop: 860.  
 Location/Qualifiers  
 1. ..871  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: PHOS1; Site\_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds M. Vaudin and B. Barrell, Oxford University Press, 1999)."

BASE COUNT 349 a 117 c 174 g 231 t  
 ORIGIN

#### Alignment Scores:

Pred. No.: 4, 83e-78 Length: 871  
 Score: 740.50 Matches: 147  
 Percent Similarity: 70.04% Conservative: 47  
 Best Local Similarity: 53.07% Mismatches: 82  
 Query Match: 22.28% Indels: 1  
 DB: 17 Gaps: 1

US-09-873-409-1 (1-659) x AZ682350 (1-871)

Qy 381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400  
 Db 39 AATAGTCTAAAGTTCGCCCATTTAATGTTTATCAACAATGATGATAGATATCCAGATAT 98  
 Qy 401 AspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPheArg 420  
 Db 99 GATTGTCAGTCTATTGGAGGTGAATGTCCTCACTGATGTAATGGAATATTAGATTGAG 158  
 Qy 421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440  
 Db 159 GATGTTCAATTTGTTTATCCACAAGACTGCTCATCATGTTATTAAGGACTTCACCTT 218  
 Qy 441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460  
 Db 219 GAAATTAAGAAAGGACAAACAATTCATTTAGTGGAGCATCAGGATGTGGGAAGTCAACT 278  
 Qy 461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 480  
 Db 279 ACTATTCAATTAATCCAAAGAAATTTATGATCCAAATGTTGGAAGAGTAACTATGACGGA 338  
 Qy 481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500  
 Db 339 AAGATATACGAGTTCGAATATCAATGTTTAAAGATCAATAGGATTAGTTGACAA 398  
 Qy 501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520  
 Db 399 GAACCAAGTGTGTTTGCAGGAACAATTCGAGAAATATTTATGCTTGGAGCTAAAGAGGA 458  
 Qy 521 ValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539

Db 459 GCAACACCACTGAAGAAGAGATGATTGAATGTGCTAAATGGCAATGCACATGACTTC 518  
 Qy 540 IleGluGlyLeuProGluLysValAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSer 559  
 Db 519 ATTTCTAACTTCAGAGAGGATATGACACATAATTCGAGAAAGAGGACATATTATCA 578  
 Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579  
 Db 579 GGAGGACAAAACAAAGAAATTCGAATTCGACGTGCAATTCGAAACCCATATTCTT 638  
 Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599  
 Db 639 CTTCTTGATGAAGCTACATCAGCACTTGATACACAAGTGAAGAAATGTACAAGACCA 698  
 Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619  
 Db 699 CTTGAAAGGATCTTAAGGAAGAACACAAATATTGTAGCACATAGACTAACCACTGTT 758  
 Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHis 639  
 Db 759 AGAAATGCAGATAAATATGTGTATTCATCAAGGAGAAATATTATGAACAAGGAACAT 818  
 Qy 640 GlnGluLeuLeuArgAsnArgAspIleTyrrPhelysLeuValAsnAlaGln 656  
 Db 819 CAAGAATTAATGGATTGCAAGGAACATATTATTGATGATTAGTCAAAAGACAA 869

RESULT 9  
 AZ670821/c 932 bp DNA linear GSS 14-DEC-2000  
 LOCUS ENTJUN69TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 DEFINITION genomic, DNA sequence.  
 ACCESSION AZ670821  
 VERSION AZ670821.1 GI:11807967  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 932)  
 Loftus, B., Van Aken, S. and Fraser, C.  
 Determination of clone end sequences from Entamoeba histolytica  
 HMI:IMSS sheared DNA library  
 Unpublished (2000)  
 CONTACT: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: b.loftus@igr.org  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
 DNA library  
 Seq primer: M13-Forward  
 Class: shotgun  
 High quality sequence start: 46  
 High quality sequence stop: 867.  
 FEATURES  
 source location/Qualifiers  
 1..932  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: pHO1; Site 1: Bet 1; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica  
 using a method described by Clark and Diamond (Clark,  
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
 method for isolate identification. Exp. Parasitol.  
 77:450 ). The DNA was mechanically sheared to give a  
 tight size distribution (~2 kb). The v + i method used for  
 the library construction is described in detail in Smith,  
 H.O. and Venter, J.C. (Making small insert libraries for  
 whole genome shotgun sequencing projects. In Genome  
 Sequencing: A Practical Approach, eds. M. Vaudin and B.

BASE COUNT 258 a 182 c 125 g 367 t  
 ORIGIN  
 Alignment Scores: 2,86e-77 Length: 932  
 Pred. No.: 734.50 Matches: 150  
 Score: 71.11% Conservative: 42  
 Percent Similarity: 55.56% Mismatches: 77  
 Best Local Similarity: 22.10% Indels: 2  
 Query Match: 17 Gaps: 1  
 DB: 1  
 US-09-873-409-1 (1-659) x AZ670821 (1-932)  
 Qy 380 TyrSerLysAlaLysSerGlyValAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399  
 Db 835 TTTGCTACAGCTAAAGCTTCTGCATATAGAAATTTATCAACAATTTAGTAGAATCCAGAT 776  
 Qy 400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419  
 Db 775 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAAATGGAAATATTATCATTA 716  
 Qy 420 ArgGluValSerPhePheTyrrProCysArgProAspValPheIleLeuArgGlyLeuSer 439  
 Db 715 GAAGATGTTCAATTTAGATATCCAAAGACCACTAAACAATTTCTTGGTGACTTGAT 656  
 Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459  
 Db 655 CTTGAAATTAAGAAGGACAAACAGTTGCAATTAGTAGGAGCATCAGGATGTGTTAAATCA 596  
 Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrrAspProValGlnGlyGlnValLeuPheAsp 479  
 Db 595 ACTACTATTCAATTAGTCCAAAGAAATTTATGATCCAGTTGGTGGTCAATTAATTAGAT 536  
 Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499  
 Db 535 GGAAAGATTTAAGAGATTTAAATATCAATGGTTAAGAAATCAATAGGATTTAGTTGA 476  
 Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrrGlyAspAsnSer 519  
 Db 475 CAAGAACCTATTGTTTGCATGTACTATTAGAGAAATATATTATCTCTGGAGCTAGAT 416  
 Qy 520 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 538  
 Db 415 GGAGAAACACCAACTGAAGAAAGAGATGTTGAATGTCTAAATGGCAATGCACATGAA 356  
 Qy 539 PheIleGluGlyLeuProGluLysTyrrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeu 558  
 Db 355 TTTATTTCTCATCTTCCAGAGGATATATACAAATGGTAGGAGAAAAGGAGCTCAATTA 296  
 Qy 559 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 578  
 Db 295 TCAGGAGGACAAAACAAAGAAATTCCTATTGCACCTGCTGATTGATTAGAAACCTACAAT 236  
 Qy 579 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 598  
 Db 235 TTATTACTTGTATGAAGCTACATCAGCACTTGATACACAAGTGAAGAAATTTGTACAA 176  
 Qy 599 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 618  
 Db 175 GCACCTGAAAAGCTAGTCAAGGAAGAACAAATTTGTTGTAGCACATAGATTAACTACT 116  
 Qy 619 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 638  
 Db 115 GTTGAATATGAAGTGAAGATTTGTTGTTTCCATCAAGGAGAAATTTATGAACAGGAACA 57  
 Qy 639 HisGlnGluLeuLeuArgAsnArgAspIle 648  
 Db 56 CATCAAGAAATTAAGGATTTAAAGAGGATTTA 27  
 RESULT 10  
 AZ683753 947 bp DNA linear GSS 14-DEC-2000  
 LOCUS AZ683753

```

DEFINITION  ENT1196TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
ACCESSION   AZ683753
VERSION     AZ683753.1  GI:11820899
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
ORGANISM    Entamoeba histolytica.
REFERENCE   1 (bases 1 to 947)
AUTHORS     Loftus,B., Van Aken,S. and Fraser,C.
TITLE       Determination of clone end sequences from Entamoeba histolytica
            HMI:IMSS sheared DNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjlloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 40
            High quality sequence stop: 778.
            Location/Qualifiers
              1..947
                /organism="Entamoeba histolytica"
                /strains="HMI:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
                /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from broth cultures of E. histolytica
                using a method described by Clark and Diamond (Clark,
                C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                method for isolate identification. Exp. Parasitol.
                77:450.). The DNA was mechanically sheared to give a
                tight size distribution (~2 kb). The v + i method used for
                the library construction is described in detail in Smith,
                H.O. and Venter, J.C. (Making small insert libraries for
                whole genome shotgun sequencing projects. In Genome
                Sequencing: A Practical Approach, eds. M. Vaudin and B.
                Barrell, Oxford University Press, 1999)."
BASE COUNT  379 a 121 c 178 g 269 t
ORIGIN
Alignment Scores:
Pred. No.:    1,558-76      Length:      947
Score:        728.50      Matches:    149
Percent Similarity: 69.82%   Conservative: 43
Best Local Similarity: 54.18% Mismatches: 81
Query Match:  21.92%      Indels:     2
DB:           17          Gaps:         1

US-09-873-409-1 (1-659) x AZ683753 (1-947)

QY  383 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleaspSer 402
Db  69 GCAAAGCATGACGCTATATAATATTATACAAATGATAGATATCCACATATTGATGT 128
QY  403 ArGSerGlnGluLysLysProAsnThrCysGluGluYasnLeuGluPheArgGluVal 422
Db  129 CAGTCTATTGGAGGTGAATGTCCTCACTGAGTGTAAATGGGAATATTAGATTGATGTT 188
QY  423 SerPhePheTyrProCysArgProaspValPheIleLeuArgGlyLeuSerLeuSerIle 442
Db  189 CAATTGTGTTATCCAAACAGACCATCTCATCATGATTATTAAGGAATTCACCTTGAAT 248
QY  443 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlySerThrSerVal 462
Db  249 AGAAGAGGAATATATTCATTAGTTGGAGCATCAGGATAGAAAGTCAACTACTATT 308

```

```

QY  463 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlnGlyGlnValLeuPheaspGlyValasp 482
Db  309 CAATTAAATCAAAGAATATTATGAACCAATGGTGGAGAGTAACATTAGATGGGAAGAC 368
QY  483 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 502
Db  369 ATAAGAGAGCTTGAATATCAATGGTTTAAAGAAATCAATAGGAATAGTTGGACAAGAACCA 428
QY  503 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly--AspAsnSerArgVal 521
Db  429 GTATTATTTTCAGGAACAATTGGAGAAATATCATCTCTGGAGCTTAAAGAGAGAAACA 488
QY  522 ValProLeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 541
Db  489 CTAAGTAAAGAGAGAGATGATTGAATGTACTAAATGGCAATGCACATGACATTTATTTCT 548
QY  542 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 561
Db  549 AAACCTCCAGAGAGGATATGACACAATAATTGGAGAAAAGGAGCATTTATTATCAGGAGGA 608
QY  562 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeu 581
Db  609 CAARAACAAGAATTCGAATTCACGTGATTCGATTCGAAACCCATCTATTCTTCTTCT 668
QY  582 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLeuValValGlnHisAlaLeuAsp 601
Db  669 GATGAAGCTACATCAGCACTTCATACACAAGTGAAGAAATGTGTACAGAAGCACTTGAC 728
QY  602 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 621
Db  729 AAAGCATCAAAAGGAGGAACAACAATTTATTGTAGGACATAGACTATCGACTATTCAAAAT 788
QY  622 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 641
Db  789 GCAGATCAATATGTGTTATTATGAGAGGAAAATAGTAGAACAAGGAACACATCAAGAA 848
QY  642 -LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 655
Db  849 TTTATTGGATTTGAAAAGATTTTATTATATACACTTGCCTTATGCA 891

RESULT 11
AZ540627
LOCUS      ENTEQ18TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION  AZ540627
VERSION    AZ540627.1  GI:11147603
KEYWORDS   GSS.
SOURCE     Entamoeba histolytica.
ORGANISM   Entamoeba histolytica.
REFERENCE  1 (bases 1 to 886)
AUTHORS    Loftus,B., Van Aken,S. and Fraser,C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
            HMI:IMSS sheared DNA library
JOURNAL    Unpublished (2000)
COMMENT    Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjlloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 23
            High quality sequence stop: 886.
            Location/Qualifiers
              1..886
                /organism="Entamoeba histolytica"

```





```
Db 423 ATGTATGCTCAGAGTTTGGAGGTACCATACAGAAATCTTTGAGGAAAGCACACATCTTT 482
Qy 321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
Db 483 GGAATTACATTTCTTCACCCAGGCAATGATGATATTTTCTCTATGCTGGATGTTTCGG 542
Qy 341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
Db 543 TTTGGAGCTACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTT 602
Qy 361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
Db 603 TCAGCTGTGTGTTGGTGCCATGCGCGGCAAGTCAAGTTCATTTGCTTCCTGACTAT 662
Qy 381 Ser-LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db 663 GCCCAAGCCAAATATACAGAGCCACATCATGATCATGATGATGATGATGATGATGAT 722
Qy 400 eAspSerArgSer-GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu---P 419
Db 723 TGACAGCTACAGCACCGGAGCCCTAATGCGCAACACATTCGGAAGGAATGTCACCATTT 782
Qy 419 heArgGluValSerPhePheTyrPro 427
Db 783 GGTGAAGTTGTATTTCAACTATCCC 808

RESULT 14
A2541090 897 bp DNA linear GSS 14-NOV-2000
LOCUS ENTDS67TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION A2541090
VERSION A2541090.1 GI:11148493
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 897)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HM1:IMSS sheared DNA library
COMMENT Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 81
High quality sequence stop: 890.
Location/Qualifiers
1..897
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1, Site 1: Bet 1: Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
```

```
BASE COUNT 335 a 110 c 177 g 275 t
ORIGIN
Alignment Scores: 1,31e-75 Length: 897
Pred. No.: 720.50 Matches: 143
Score: 74.59% Conservative: 39
Best Local Similarity: 58.61% Mismatches: 61
Query Match: 21.68% Indels: 1
DB: 17 Gaps: 1
US-09-873-409-1 (1-859) x AZ541090 (1-897)
Qy 414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 433
Db 94 AAAGTGAATTTGAATTTAAAGACATTTGTTTCAGATATCCAAACAGACCAGACAAATCT 153
Qy 434 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
Db 154 GTCTTGAAGGTATTTCAATTCAAAGTAGAACAGAAAAAACTGTTGCAATTAGTAGGAGCA 213
Qy 454 SerCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 473
Db 214 TCAGGATGTGGTAAATCAACATCAGTTGATTGAAAGATTTTATGATCCACACAT 273
Qy 474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
Db 274 GGAGATGATTTATGATGACATAATATCAAGATTTGAATATTCATTTCTTAAGAAGT 333
Qy 494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
Db 334 CAAATTGGAATGGTAGGACAAAGAACCAAGTATTTGCTGAAAGTGTATGCGATAATATT 393
Qy 514 AlaTyrGlyAspAsnSerArgVal---ValProLeuAspGluIleLysGluAlaAlaAsn 532
Db 394 AGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAACAAAATTTATGCTGCTGCTAAA 453
Qy 533 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 552
Db 454 ATGCAATGCAATGACTTTATTTTCAGCAATGCCAGAGGATATACACAATGGTAGGT 513
Qy 553 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 572
Db 514 GATAGAGTGCACAAATTTCCAGGAGGACAAACAAAGAAATGCTATTGCGACGTCATTG 573
Qy 573 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 592
Db 574 ATTGAATCCAAAAGTGTATTATTCTCGATGAAGCTACATCAGCACTTGATTCAGAAAGT 633
Qy 593 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 612
Db 634 GAAAGATTTGACCAAGATGCTGACCAAGCAACCAAGGAAAGGAAAGCAATTTGTAATT 693
Qy 613 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 632
Db 694 GCATAGATTATCAACTATTTCAAAATGCAGATCAAAATATGTTATTATCAGAGGAAGA 753
Qy 633 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhenylsLeu 652
Db 754 ATTGAGAAAGAGGAACACATCAAGAGTTATTAGATTGGAAGGATTTTATTATACACTT 813
Qy 653 ValAsnAlaGln 656
Db 814 GCTATGCACAA 825
RESULT 15
BH155700
LOCUS ENTRO54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION BH155700
VERSION BH155700.1 GI:15727822
```





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 96.5072 Seconds  
(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-1

Perfect score: 3323  
Sequence: 1 MIAEKGAHELMKRLGLYS.....QELLRNDRIFKLVNAQSVQ 659

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop .6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_epool/US09873409/runat 27032003 115422 19259/app\_query.fasta\_1.7544  
-DB=issued Patents NA -QMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409@cgn\_1 1.578 @runat 27032003 115422 19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGOQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1872	56.3	2726	1 US-08-461-823-1	Sequence 1, Appli
2	1872	56.3	4264	2 US-08-784-649A-1	Sequence 1, Appli
3	1872	56.3	4264	2 US-08-784-649A-5	Sequence 5, Appli
4	1872	56.3	4646	1 US-08-181-471-2	Sequence 2, Appli
5	1872	56.3	4669	6 5206352-3	Patent No. 5206352
6	1872	56.3	6505	2 US-08-793-610-5	Sequence 5, Appli
7	1872	56.3	9318	2 US-08-793-610-6	Sequence 6, Appli
8	1855	55.8	4669	2 US-08-752-447-1	Sequence 1, Appli
9	1855	55.8	4669	4 US-09-316-167-1	Sequence 1, Appli
10	1849.5	55.7	4233	3 US-09-120-513-1	Sequence 1, Appli
11	1849.5	55.7	4233	4 US-09-450-105-1	Sequence 1, Appli
12	1845	55.5	4669	2 US-08-583-276-18	Sequence 18, Appl

13	1254	37.7	4002	2	US-08-996-545-1	Sequence 1, Appli
14	1254	37.7	4002	2	US-08-996-545-3	Sequence 3, Appli
15	1254	37.7	4002	4	US-09-328-320-1	Sequence 1, Appli
16	1254	37.7	4002	4	US-09-328-320-3	Sequence 3, Appli
17	1227.5	36.9	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1184.5	35.6	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1151	34.6	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	1049.5	31.6	4224	1	US-08-612-521-1	Sequence 1, Appli
21	975	29.3	3909	3	US-08-232-537-1	Sequence 3, Appli
22	851	25.6	6143	1	US-08-612-521-3	Sequence 3, Appli
23	842	25.3	3924	2	US-08-996-644-3	Sequence 3, Appli
24	842	25.3	3924	3	US-09-352-552-3	Sequence 3, Appli
25	842	25.3	3927	2	US-08-996-644-1	Sequence 1, Appli
26	842	25.3	3927	3	US-09-352-552-1	Sequence 1, Appli
27	840	25.3	3792	4	US-09-351-224E-10	Sequence 10, Appli
28	783	23.6	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	677	20.4	2376	1	US-08-394-880B-1	Sequence 1, Appli
30	644	19.4	2244	4	US-09-061-764A-18	Sequence 18, Appli
31	623.5	18.8	2061	4	US-09-061-764A-17	Sequence 17, Appli
32	608.5	18.3	5120	3	US-08-772-270A-6	Sequence 6, Appli
33	608.5	18.3	8370	2	US-08-488-706-1	Sequence 1, Appli
34	601	18.1	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
35	600.5	18.1	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
36	599	18.0	13188	4	US-08-961-527-70	Sequence 70, Appli
37	594.5	17.9	7721	3	US-08-772-270A-14	Sequence 14, Appli
38	555.5	16.7	1959	4	US-09-061-764A-4	Sequence 4, Appli
39	555.5	16.7	7760	4	US-08-961-527-63	Sequence 63, Appli
C 40	552.5	16.6	28804	2	US-08-592-874-1	Sequence 1, Appli
C 41	552.5	16.6	28804	3	US-09-096-942-2	Sequence 2, Appli
C 42	552.5	16.6	28804	3	US-09-096-867-2	Sequence 2, Appli
43	548	16.5	6492	4	US-08-961-527-188	Sequence 188, Appli
44	544.5	16.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
45	543.5	16.4	7186	4	US-08-961-527-39	Sequence 39, Appli

ALIGNMENTS

RESULT 1

US-08-461-823-1  
; Sequence 1, Application US/08461823  
; Patent No. 5593840  
; GENERAL INFORMATION:  
; APPLICANT: Bhatnagar, Satish K.  
; APPLICANT: George Jr., Albert L.  
; APPLICANT: Nazarenko, Irina  
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OncorPharm, Inc.  
; STREET: 200 Perry Parkway  
; CITY: Gaithersburg  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20877  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461.823  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/168,621  
; FILING DATE: 16-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/010,433  
; FILING DATE: 27-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Karta, Glenn E.  
; REGISTRATION NUMBER: 30,649  
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 301 527-2058  
 TELEFAX: 301 208-6997  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2726 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-461-823-1

Alignment Scores:  
 Pred. No.: 2,8e-219 Length: 2726  
 Score: 1872.00 Matches: 356  
 Percent Similarity: 74.22% Conservative: 142  
 Best Local Similarity: 53.06% Mismatches: 159  
 Query Match: 56.33% Indels: 14  
 DB: 1 Gaps: 4

US-09-873-409-1 (1-659) x US-08-461-823-1 (1-2726)

Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
 Db 317 GTCAATTGTGAGAAAGGAATCATGATGAACCTCATGAAAGAGAGGCACTTACTTCAA 376  
 Qy 21 LeuValMetSerGln-----AspIleLysAlaAspGluMet 34  
 Db 377 CTTGTCACAATCGACAGCAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 436  
 Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
 Db 437 AAAAGTGAATGATGCGCTTGGAAATGCTTCAATGATTCAGATCCAGTCAATAAGA 496  
 Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluSer 67  
 Db 497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAAGCCCAAGCAAGCTTAGTACC 556  
 Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 87  
 Db 557 AAAGAGGCTCTGGATGAATGATACCTCCAGTTCTCTTTGGAGGATTAATGAAGCTAAAT 616  
 Qy 88 LysProGluTyrProValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
 Db 617 TTAACGTAATGCGCTTATTTGTGTGGTGTATTTGTGCGCATATATAATGGAGCCCTG 676  
 Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126  
 Db 677 CAACCGAGCATTTGCAATAATATTTTCAAGATATATAGGGGTTTTTACAAGAATTGATGAT 736  
 Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146  
 Db 737 CCGTAAACAAACACGACAGATAATGTAATCTTCTACTATTGTTCTAGCCCTTGGAAAT 796  
 Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166  
 Db 797 ATTTCTTTTATACATTTTCTTCCCTTGGGTTTCACATTTTGGCAAGCTGGAGATCTCTC 856  
 Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186  
 Db 857 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGATGTTGTTT 916  
 Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
 Db 917 GATGACCTTAAACACCACTCGAGCATTTGACTTACCAGGCTCGCCCAATGATGCTCTCAA 976  
 Qy 207 IleGlnGlyValThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
 Db 977 GTTAAAGGGGCTATAGTTTCCAGGCTTGTGTAATTTACCAGAAATATAGCAATCTTGG 1036  
 Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246

Db 1037 ACAGGAATAATATATATCTCTTATGTTGGCAACTAAACACTGTTACTCTTACCAATT 1096  
 Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
 Db 1097 GTACCATCATTTGCAATAGCAGGAGTGTGTTGAATGAAATGTTGTCTGGACAGCACTG 1156  
 Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
 Db 1157 AAAGATAAGAAAGAACTAGAGGTCTCGGAAGATCGCTACTTGAAGCAATAGAAAACTTC 1216  
 Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
 Db 1217 CGAACCGTTGTTCTTCTGACTCAGCAGCAGAAAGTTTGAACATATATGCTCAGAGTTTG 1276  
 Qy 307 GlnThrGlnHisAspAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
 Db 1277 CAGGTACCATCAGAAACTCTTTGAGGAAAGCACACATCTTTGGAAATTACATTTTCCTTC 1336  
 Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346  
 Db 1337 ACCCAGGCAATGATGATTTTCTTATGCTGGATGTTTCGGTTTGGAGCCTACTTGGTG 1396  
 Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
 Db 1397 GCACATAAACTCATGAGCTTTCAGGATGTTCTGTAGTATTTTCAGCTGTTGCTTTGGT 1456  
 Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
 Db 1457 GCATGGCGGTGGGCAAGTCAGTTTCATTTGCTCTGATATGCCAAAGCCAAATATCA 1516  
 Qy 387 AlaAlaHisLeuPheAlaLeuLeuLysLysLysProAsnIleAspSerArgSerGlnGlu 406  
 Db 1517 GCAGCCCATCATCATCATCATCATGATCATTTGAAACCCCTTTGATTGACAGCTACAGCGAA 1576  
 Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426  
 Db 1577 GCGCTTAATGCGGAAACACATTTGGAAGAAATGTACATTTGTTGTAAGTTGTATTCAACTAT 1636  
 Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
 Db 1637 CCCACCCAGCGGACATCCAGTGTCTCAGGACCTGAGCCTGGAGTGAAGAGGCGCAG 1696  
 Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
 Db 1697 AGCTGTGCTCTGTTGGGCGAGCAGTGGCTGTGGGAAGACACAGTGTCCAGCTCTCGAG 1756  
 Qy 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
 Db 1757 CGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGTATGGCAAGAAATAAAGCGACTG 1816  
 Qy 487 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
 Db 1817 AATGTTCTAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCTGTTTAC 1876  
 Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526  
 Db 1877 TGCAGCATTTGCTAGAAACATTTGCCCTATGGAGACACACGCGGCTGTGTCCAGGAGAG 1936  
 Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
 Db 1937 ATCTGTGAGGCGAGCAAGAGGCGCAACATACATGCTTTCATCGAGTCACTGCTTAATAA 1996  
 Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
 Db 1997 TATAGCATTAAGTAGGAGACAAAGAACTCAGCTCTCTGTTGGCCAGAAACAGCACTT 2056  
 Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586  
 Db 2057 GCATAGCTCGTGGCTTGTAGACAGCTCATATTTTGTCTTTTGTGATGAGCCAGCTCA 2116  
 Qy 587 AlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606

```
Db 2117 GCTCTGGATACAGAAAGTGAAGGTTGTCTCAAGAGCCCTGGACAAAGCCAGAGAGGC 2176
Qy 607 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 2177 CGCACCTGCAATGTGTGCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTG 2236
Qy 627 ValLeuHisAsnGlyLysIleGlyGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 646
Db 2237 GTGTTTCAGAAATGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGCTGCTGGCACAGAAA 2296
Qy 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db 2297 GGCATCTATTTTCAATGCTCAGTGTCCAGGCT 2329

RESULT 2
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-784-649A-1

Alignment Scores:
Pred. No.: 6,17e-219 Length: 4264
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 2 Gaps: 4

US-09-873-409-1 (1-659) x US-08-784-649A-1 (1-4264)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 1951 GTCATTGTGGAGAAAGAAATCATGTAACATCATGAAAGAGAAAGGCACTTACTTCAAA 2010
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 2011 CTTGTGCAATGCACAGCAGGAAATGAAGTTGAATTAGAAATGTCAGCTGATGAATCC 2070
```

```
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2071 AAAAGTGAATTCATGCTTGGAAATGCTCTCAATGATTCACAGATCCAGTCTAATAAGA 2130
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 2131 AAAAGATCAACTCGTAGAGTGTCCTGGATGCACAGCCCAAGACAGAAAGCTTAGTACC 2190
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2191 AAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCTTTTGGAGGATATATGAAGCTAAAT 2250
Qy 88 LysProGluIlePheProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2251 TTAACCTGAATGGCCTTATTTTGTGGTGTATTTTGTGCATTATATAAATGGAGGCGTG 2310
Qy 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2311 CAACGAGCAATTCGAATATATTTTCAAGATATATAGGGGTTTTTACAGAAATGTGATAT 2370
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2371 CCTGAACAAACACGACAGAAATAGTAACTTGTTTCACTATTGTTTCTAGCCCTTGGAAAT 2430
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2431 ATTTCTTTTATATACATTTTCTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 2490
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 186
Db 2491 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGAGATGTGATGGTTT 2550
Qy 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 2551 GATCACCTTAAACACACCATGGAGCATTGACTACAGGCTCGCAATGATGCTGCTCAA 2610
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2611 GTTAAAGGGCTATAGGTTCCAGGCTTGCTGTAATACCAGAAATATATAGCAAAATCTTGG 2670
Qy 227 LeuSerValIleIleSerPheIleTyrGlyThrGluMetThrPheLeuIleLeuSerIle 246
Db 2671 ACAGAAATATATATCTTCATCTATGTTGGCAATACACACTGTTACTCTTAGCAAT 2730
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 2731 GTACCCATCATGCAATAGCAGAGAGTTGTTGAATGAAATGTTGCTGGACAGCACTG 2790
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 2791 AAGATAAGAAAGAACTAGAAAGGTGCTGGAGAGATCGCTACTGAAGCAATAGAAAACCTC 2850
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 2851 CGAAACCGTTGTTTCTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTG 2910
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 2911 CAGGTACCATACAGAAACTCTTTCAGGAAAGACACACATCTTTGGAAATTACATTTCCTTC 2970
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIle 346
Db 2971 ACCAGGCAATGATGATTTTCTATGCTGATGTTTCCGGTTTCCGGTTTGGAGCCTACTTGGT 3030
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3031 GCACATAAATCATGAGCTTTGAGAGATGTTCTGTAGTATTTTTCAGCTGTTGCTCTTGGT 3090
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3091 CCATGGCCGGGGGCAAGTCAGTTCATTTGCTCTGACTATGCAAAAGCCAAATATCA 3150
```

```
QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLeuProAsnIleAspSerArgSerGlnGlu 406
Db 3151 GCAGCCACATCATCATGATCAITGAAAAACCCCTTTGATTGACAGCTACACAGCGAA 3210
QY 407 GlyLysLeuProAspThrCysGluGlyAsnLeuLeuPheArgGluValSerPheThr 426
Db 3211 GGCTTAATGCGCAACACATTTGGAAGGAATGTCACATTTGGTGAAGTTGATTCAACTAT 3270
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3271 CCACCGCAGCGACATCCAGTCTTCAGGACTGAGCCCTGGAGTGGAAGAAGGCCAG 3330
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 3331 ACCTGCTCTGCTGGCAGCAGTGGCTGGGGAAGAGCACAGTGTCTCCAGCTCTCTGGAG 3390
QY 467 ArgLeuThrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3391 CGTTTCTACACCCCTTGGCAGGGAAGTGTCTGATGGCAAGAAATAAAGCGACTG 3450
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3451 AATGTCAGTGGCTCGAGCACCTGGGATCGTGTCCAGGAGCCCATCTCTGTTTGAC 3510
QY 507 CysSerIleAlaGluAsnIleAlaTyArgAspAsnSerArgValValProLeuAspGlu 526
Db 3511 TGCAGCATTTGCTGAGAACATTTGCTATGGAGACACAGCCGGTGTGTACAGGAAGAG 3570
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 3571 ATTGTCAGGCGACCAAGGAGGCGCAACATACATGCTTCATCGAGTCACTGCGCTAATAA 3630
QY 547 TyrAsnTrpGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeu 566
Db 3631 TATAGCACTAAGTAGAGCAAGAAAGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAG 3690
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 3691 GCCATAGCTGCTGCCCTTGTAGACAGCCTCATATTTTGTGTTGGATGAAGCCACGTCA 3750
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 3751 GCTCTGGATACAGAAAGTGAAGAGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGG 3810
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 3811 CGCACTGCAATTTGATGCTCACCCGCTGTCCACCATCCAGATCCAGATCTTAATAGTG 3870
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 3871 GTGTTTCAGAAATGGCAGAGTCAAGGAGCATGGCAGCATTCAGCAGCTGTGGCAGAGAA 3930
QY 647 AspIleTyPheLysLeuValAsnAlaGlnSer 657
Db 3931 GGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 3963
```

## RESULT 3

US-08-784-649A-5

; Sequence 5, Application US/08784649A

; Patent No. 5830697

; GENERAL INFORMATION:

; APPLICANT: Sikic, Branimir I

; APPLICANT: Chen, Gang

; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO

; NUMBER OF INVENTION: CYCLOSPORIN MODULATION

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Fish &amp; Richardson

; STREET: 2200 Sand Hill Road

; CITY: Menlo Park

; STATE: CA

; COUNTRY: USA

; ZIP: 94025

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
```

```
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
```

```
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-5
```

```
Alignment Scores:
Pred. No.: 6,17e-219 Length: 4264
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 2 Gaps: 4
```

US-09-873-409-1 (1-659) x US-08-784-649A-5 (1-4264)

```
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTrSer 20
Db 1951 GTCAATTGTGGAGAAAGAAATCATGATGAACATCATGAAAGAGAGAGGCAATTTACTTCAA 2010
QY 21 LeuValMetSerGln-----AspIleLysIleAlaAspGluGlnMet 34
Db 2011 CTTGTCACAATCAGACAGACAGGAAATGAAGTTGAATAGAAATGCACGTGATGAATCC 2070
QY 35 GluSerMetThrTySerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2071 AAAAGTGAATTTGATGGCTTGGAAATGCTTCAATATGATTCAGATCCAGCTTAATAAGA 2130
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 2131 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2190
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 87
Db 2191 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTGGAGGATTAAGAGCTAAAT 2250
QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2251 TTAACGTGAATGGCCTTATTTTGTGTGTATTTTGTGGCCATTAATAATGGAGCCCTG 2310
QY 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2311 CAACAGCATTTTGAATTAATATTTTCAAGATTAATAGGGGTTTACAGAAATTTGATGAT 2370
QY 127 LysThrThrLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyVal 146
Db 2371 CCTGAACAAAAACAGACAGAAATAGTAATCTTCTTCACTATTCTTCTAGCCCTTGAAT 2430
QY 147 IleCysPheValSerTyPheMetGlnGlyLeuPhePheTyGlyArgAlaGlyGluIleLeu 166
Db 2431 ATTCTTTTATTACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2490
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaTrpPhe 186
```

Dbb 2491 ACCAAGCGGCTCCGATACATGATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2550  
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
Dbb 2551 GATGACCTTAAACACACACCTGAGCAATGACTACAGGCTCGCCAAATGATGCTCTCAA 2610  
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
Dbb 2611 GTTAAAGGGCTATAGTTTCCAGGCTTGCTGTAATATACCAATATATAGCAATCTTGGG 2670  
Qy 227 LeuSerValIleIleSerPheIleTyTrpGluMetThrPheLeuIleLeuSerIle 246  
Dbb 2671 ACAGAAATATATATCTTCACTATGTTGGCAACTAACACTGTTACTCTTACCAAT 2730  
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
Dbb 2731 GTACCCATCATTCATAGCAGGATGTTGAAATGAAATGTTGCTGGACAAGCAGCTG 2790  
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
Dbb 2791 AAGATAAGAAAGACTAGAGGCTGCGGAAGATCGCTACTGAAGCAATAGAAAACCTTC 2850  
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGluMetTyTrpGluMetLeu 306  
Dbb 2851 CGAACCGTGTGTTCTTGAAGCTCAGGAGCAGAAAGTTGAACATATATGATGCTCAGAGTTG 2910  
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyTrpAlaPhe 326  
Dbb 2911 CAGGTACCATACAGAAACTCTTTGAGGAAGCACAATCTTTGGAAATACATTTTCCCTTC 2970  
Qy 327 SerHisAlaPheIleTyTrpPheAlaTyTrpAlaAlaGlyPheArgPheGlyAlaTyTrpLeu 346  
Dbb 2971 ACCCAGGCAATGATGATTTTCTATGCTGATGTTCCGGTGTGGAGCCTTACTTGGTG 3030  
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyGly 366  
Dbb 3031 GCACATAAATCACTGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTTCTCTTGTG 3090  
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyTrpSerLysAlaLysSerGly 386  
Dbb 3091 GCATGCGCGTGGGCAAGTCAGTTCATTTGCTCTGACTATGCCAAGCCAAATATCA 3150  
Qy 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
Dbb 3151 CGAGCCCATCATCATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3210  
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyTr 426  
Dbb 3211 GGCCTAATGCCCAACACATTTGGAAGGAATGTGCATTTTGGTGAAGTGTGATTTCAACTAT 3270  
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
Dbb 3271 CCCACCGACCGACATCCAGTGTCTTACGAGCTGAGCTGAGGTGAAGAGGCGCAG 3330  
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
Dbb 3331 ACCTGGCTCTGTGGGACAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAG 3390  
Qy 467 ArgLeuTyAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
Dbb 3391 CGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGTATGTCGCAAGAAATAAAGGCGACTG 3450  
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
Dbb 3451 AATGTTCACTGGCTCGAGACACCTGGGCACTGCTTCCAGAGCCCACTCTGTTGAC 3510  
Qy 507 CysSerIleAlaGluAsnIleAlaTyTrpGlyAspAsnSerArgValValProLeuAspGlu 526  
Dbb 3511 TGCAGCATTTGCTGAGAACATTTGCTATGAGACACACAGCGGGTGTGTGCACAGGAAGAG 3570  
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
Dbb 3571 AATGTGAGGGCGCAAGAGGCGCAACATACATACCTTCTGAGTCACTGCTGCTAATAAA 3630

Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
Dbb 3631 TATAGCACTAAGTAGGACAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACACGCAAT 3690  
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586  
Dbb 3691 GCCATAGCTGTGGCTTGTAGACAGCTCATATTTTGTCTTTTGGATGAAGCCACGTC 3750  
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
Dbb 3751 GCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGGC 3810  
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
Dbb 3811 CGCAGCTGATTTGATTTGCTCAGCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 3870  
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646  
Dbb 3871 GTGTTTTCAGATGCGCAGTCAAGGAGCATGGCAGCATGAGCTGCTGGCACAGAA 3930  
Qy 647 AspIleTyTrpPheLysLeuValAsnAlaGlnSer 657  
Dbb 3931 GCATCTATTTTCAATGCTCAGTGTCCAGGCT 3963  
RESULT 4  
US-08-181-471-2  
; Sequence 2, Application US/08181471  
; Patent No. 5641508  
; GENERAL INFORMATION:  
; APPLICANT: Li, Lingna  
; APPLICANT: Lishko, Valeryi K.  
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Thomas Fitting  
; STREET: 12526 High Bluff Drive, Suite 300  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92130  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/181,471  
; FILING DATE: 13-JAN-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/041,553  
; FILING DATE: 02-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: ANT0029P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-792-3680  
; TELEFAX: 619-792-8477  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4646 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 425..4267

US-08-181-471-2

## Alignment Scores:

Pred. No.: 7,18e-219 Length: 4646  
Score: 1872.00 Matches: 356  
Percent Similarity: 74.22% Conservative: 142  
Best Local Similarity: 53.06% Mismatches: 159  
Query Match: 56.33% Indels: 14  
DB: 1 Gaps: 4

US-09-873-409-1 (1-659) x US-08-181-471-2 (1-4646)

QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluMetAlaLysArgGlyLeuTyrSer 20  
DB 2237 GTCAATTTGGAGAAAGGAATCATGATGTAAGTCTCATGAAAGAGGCAATTTACTTCAAA 2296  
QY 21 LeuValMetSerGln-----LysSerIleLys-----AspIleLysLysAlaAspGluGlnMet 34  
DB 2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTTAGAAATGCAGCTGAATGAATCC 2356  
QY 35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
DB 2357 AAAAGTGAAATTCATGCTTGGAAATGCTCTCAATGATTCAGATCCAGCTCAATAAGA 2416  
QY 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluSer 67  
DB 2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAAGCCCAAGACAGCAAGCTTAGTACC 2476  
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
DB 2477 AAGAGGCTCTGGATGAAGATATACCTCCAGTTCTCTTTGGAGGATATGAAGCTAAAT 2536  
QY 88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
DB 2537 TTAACCTGAATGGCTTATTTGTGTGGTGATTTGTGGCCATTATTAATGAGGCGCTG 2596  
QY 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126  
DB 2597 CAACCAGCATTTTGCATATATATTTTCAAGATATATAGGGGTTTTTACAAGAAATGATGAT 2656  
QY 127 LysThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146  
DB 2657 CCTGAAACAAACAGCAGAAATAGTACTTGTGTTTCTACTATGTTTCTAGCCCTTGGAAAT 2716  
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166  
DB 2717 ATTTCTTTTATTACATTTTCTTCAGGTTTCCACATTTGGCAAGCTGGAGAGATCCTC 2776  
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186  
DB 2777 ACCAAGGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836  
QY 187 AspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
DB 2837 GATGACCCCTAAAAACACCATCGAGCATGATGACACCAGGCTCGCCCAATGATGCTGCTCAA 2896  
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
DB 2897 GTTAAAGGGGCTATAGTTCAGGCTTGTGTAATTTACCAGATATAGCAATCTTGGG 2956  
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246  
DB 2957 ACAGGAATAATTATATCTTCATCTATGTTGGCAACTAACACTGTCTTACTCTTAGCAATT 3016  
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
DB 3017 GTACCCATCATTCATAGCAGAGTGTGTAATTTGAAATGAAATGTGTCTGGACAAAGCAGCTG 3076  
QY 267 LysAspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
DB 3077 AAAGATAAGAAAGAACTAGAGGTGTGGAGAGATCGCTACTAGCAACATAGAAACTTC 3136  
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306

DB 3137 CGAACCCTGTTGTTCTTCTGACTCAGGACGAGAAGTTTGAACATATGTATCTCAGAGTTTG 3196  
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
DB 3197 CAGGTACCATACAGAACTCTTTTGGAGAAAGCACATCTTTGGAAATTACATTTTCTTCTC 3256  
QY 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346  
DB 3257 ACCCAGGCAATGATGATTTTCTTATGCTGATGTTCCGGTTTGGAGCTTACTTGGTG 3316  
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
DB 3317 GCACATAAATCATGAGCTTTCAGAGATGTTCTGTAGTATTTTTCAGCTGTGTCTTGTGT 3376  
QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
DB 3377 GCATGGCGCTGGGCAAGTCACTTTCATTTGCTCTGACTATGCCAAGCCAAATATATCA 3436  
QY 387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
DB 3437 CGAGCCCATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGACGGAA 3496  
QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426  
DB 3497 GGCCTAATGCCAACAACATTTGAAGGAAATGTACATTTTGGTGAAGTTGTATTCAACTAT 3556  
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
DB 3557 CCACCCGACCCGAGATCCAGTGCTTTCAGGAGCTGAGCTGGAGTGAAGAGGCCAG 3616  
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 466  
DB 3617 ACGTGGCTCTGTGGCAGCAGTGGCTGTGGAAAGACACAGTGTCTCCAGCTCTCTGGAG 3676  
QY 467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
DB 3677 CGGTTCTAGACCCCTTGGCAGGAAAGTGTCTGTGTCGCAAGAAATAAAGGACCTG 3736  
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
DB 3737 AATGTCAGTGGCTCGAGCACACCTGGCATCTGTCTCCAGAGCCCATCTCTGTTTAC 3796  
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526  
DB 3797 TGCAGCATTTGCTGAGAACATTCCTTATGGAGACAAACAGCCGGTGGTGTACAGAGAG 3856  
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
DB 3857 ATCGTGAGGGCGCAAGAGGCCAACATACATGCTTTCATCGAGTCACTGCTTAATAA 3916  
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
DB 3917 TATAGCATTTAAAGTAGGAGACAAAGAACTCAGCTCTCTGTGTGGCCAGAAACAGCAT 3976  
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 586  
DB 3977 GCCATAGCTCGTGGCTTGTAGACAGCTCATATTTTGTCTTTGGATGAGCCAGCTCA 4036  
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
DB 4037 GCTCTGGATACAGAAAGTGAAGGTTGTCTCAAGAAAGCCCTCGCAAAAGCCAGAGAGC 4096  
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
DB 4097 CCACCTGATGATGATGCTCCCGCTGTCCACATCCAGATCCAGACTTAATAGTG 4156  
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg 646  
DB 4157 GTGTTTCAGAAATGGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGTGTCACAGAAA 4216  
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657



Db 4217 GGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

```
RESULT 5
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3:
; LENGTH: 4669
5206352-3
```

```
Alignment Scores:
Pred. No.: 7,24e-219 Length: 4669
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: 6 Gaps: 4
```

US-09-873-409-1 (1-659) x 5206352-3 (1-4669)

```
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSer 20
Db 2237 GTCATTTGGAGAAAGGAATCATGATGAATCATGAAAGAAAGGCAATTTCTTCAA 2296
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 2297 CTGTGCACAAATCAGACAGCAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 2356
Qy 35 GluSerMetThrTyr-SerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db 2357 AAAAGTGAAATTTGATGCCTTGGAAATGTCTTCAATGATTCAGATCCAGCTCTAATAAGA 2416
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAGAGCCACAGACAGAAAGCTTAGTACC 2476
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 2477 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTTATGAAGCTAAT 2536
Qy 88 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 2537 TTAACCTGAATGACCTTATTTGTGTGTGTGATTTTGTGCCATATATAATGGAGCCCTG 2596
Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 2597 CAACAGCAATTTGCATATATATTTTCAAGATATAGGGGTTTTTACAGAAATGATGAT 2656
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 2657 CCTGAAACAAAACGACAGAAATAGTAATCTTGTTCCTACTATTGTTCTAGCCCTTGGAAAT 2716
Qy 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 2717 ATTTCTTTTATTACATTTTCTTCCAGGTTTTCACATTTTGGCAAGCTGGAGAGATCTCTC 2776
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrClnAspIleAlaTyrPhe 186
Db 2777 ACCAAGCGGCTCCGATACATGTTTTCGATCCATCTCAGACAGGATGTGAGTTGGTTT 2836
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
```

```
Db 2837 GATGACCTAAAAACACACATCGGAGCATTTGACTACAGGCTCGCCAAATGATGCTGCTCAA 2896
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 2897 GTTAAAGGGCTATAGTTCCAGGCTTGCTGTATATACCCAGAAATATAGCAAACTCTGGG 2956
Qy 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerIle 246
Db 2957 ACAGGAATAATTATATCTTCTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 3017 GTACCCCATCTTGAATAGCAGGAGTTGTTGAATGAAATGTTGCTGGCAACAGCATG 3076
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 3077 AAGATAGAAGAACTAGAAAGTGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 3137 CGAACCGTTGTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTG 3196
Qy 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 3197 CAGGTACCATACAGAACTCTTTGAGGAAAGCAGACATCTTTGGAAATTTACATTTTCCCTTC 3256
Qy 327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValatYrLeuIle 346
Db 3257 ACCCAGGCAATGATGATTTTCTATGCTGGATGTTCCGGTTTGGAGCTTACTTGGTG 3316
Qy 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 3317 GCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTTCTTGGT 3376
Qy 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 3377 GCCATGGCCGTGGGCAAGTCAGTTTCATTTGCTCTGATATGCCAAAGCCCAAAATATCA 3436
Qy 387 AlaAlaHisLeuPheAlaLeuLeuLysLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 3437 CGAGCCCAATCATCATGATCATTTGAAAGAAACCCCTTTGATTTGACAGCTCAGCAGCGAA 3496
Qy 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db 3497 GGCCTAATGCGAACACACATTCGAAGGAAATGTTCACATTTTGGTGAAGTTGTATTCAACTAT 3556
Qy 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 3557 CCCACCCGACCGACATCCAGTGTCTCAGGAGCTGAGCCCTGGAGGTGAAGAGGCGCAG 3616
Qy 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db 3617 AGCTGGCTCTGGTGGGCGAGCAGTGGCTGTGGGAAGACACACAGTGTCTCAGCTCTCGAG 3676
Qy 467 ArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 3677 CGGTTCTACGACCCCTTCGCGAGGAAAGTGTCTGTGATGGCAAGAAATAAAGCGCATG 3736
Qy 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 3737 AATGTTCTAGTGGCTCCGAGCACCATCTGGGCATCTGTGTCCAGGAGCCCATCTGTTTAC 3796
Qy 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db 3797 TGCACATTTGCTGAGAAATTCCTGATGAGACAAACACCGCGGTGTGTACAGGAAGAG 3856
Qy 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlyLys 546
Db 3857 ATCTGAGGGCGAGCAAGAGGCGCAACATACATGCTTCATCGACTCAGTGCCTAATAAA 3916
Qy 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
```

Db 3917 TATAGCACTAAAGTAGGAGACAAGGAAGTCTCTGCTGGCCAGAAAACAGCAT 3976  
Qy 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586  
Db 3977 GCCATAGCTCGTCCCTTGTGTAGACAGCTCATATTTTGTCTTTGGATGAAGCCAGTCA 4036  
Qy 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
Db 4037 GCCTGTGATACAGAAGAGTGAAGAGTTGTCCAAAGAGCCCTGGACAAAACCCAGAGAGC 4096  
Qy 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
Db 4097 CGCACTGCAATGTGATGTCTACCGCTGTCCACCATCCAGATCCAGATTAATAGTG 4156  
Qy 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646  
Db 4157 GTGTTTCAGAAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4216  
Qy 647 AspileTyRPhelyLeuValAsnAlaGlnSer 657  
Db 4217 GGCATCTATTTTCAATGGTCACTGCTCCAGGCT 4249  
RESULT 6  
US-08-793-610-5  
; Sequence 5, Application US/08793610  
; Patent No. 5858744  
; GENERAL INFORMATION:  
; APPLICANT: BAUM, Christopher  
; APPLICANT: STOCKING-HARBERS, Carol  
; APPLICANT: OSTERTAG, Wolfram  
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,610  
; FILING DATE: 07-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 31 973.8  
; FILING DATE: 08-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 03 952.1  
; FILING DATE: 07-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/03175  
; FILING DATE: 10-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berman, Richard J.  
; REGISTRATION NUMBER: 39,105  
; REFERENCE/DOCKET NUMBER: P1614-7007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6505 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA  
US-08-793-610-5

Alignment Scores: 1.3e-218 Length: 6505  
Pred. No.: 1872.00 Matches: 356  
Score: 74.22% Conservative: 142  
Best Local Similarity: 53.06% Mismatches: 159  
Query Match: 56.33% Indels: 14  
DB: 2 Gaps: 4  
US-09-873-409-1 (1-659) x US-08-793-610-5 (1-6505)  
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrrSer 20  
Db 3629 GTCATTTGGAGAAAGAAATCATGATGAACATCATGAAGAGAAAGGCAATTTCTCTCAA 3688  
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34  
Db 3689 CTTGTCACAATGCACAGCAGCAGGAATGAAGTTGAATTAGAAAAATGCAGCTGATCAATCC 3748  
Qy 35 GluSerMetThrTyrrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53  
Db 3749 AAAAGTGAATTTGATGCCTTGGAAATGTCTCAATATGATTCAGATCCAGTCTAATAAGA 3808  
Qy 54 -----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67  
Db 3809 AAAAGATCAACTGTAGGAGTGTCCGTGATCACAAAGCCCAAGACAGAAAGCTTAGTACC 3868  
Qy 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
Db 3869 AAAGAGGCTCTGGAGAAAGTATACCTCCAGTTCTTTTGGAGGATTTATGAAGCTAAAT 3928  
Qy 88 LysProGluTyrPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
Db 3929 TTAAGTGAATGGCCTTATTTTGTGTGTATTTTGTGTCATTTATTAATGGAGGCTG 3988  
Qy 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126  
Db 3989 CAACAGCATTTGCAATATATTTTCAAGATTTATAGGGGTTTTTACAGAAATTCATGAT 4048  
Qy 127 LysThrThrLeuLysHisAspAlaGluIleTyrrSerMetIlePheValIleLeuGlyVal 146  
Db 4049 CCTGAACAAACAGACAGATAGTAACTTGTTTTCACATTTGTTTCTAGCCCTTGAAT 4108  
Qy 147 IleCysPheValSerTyrrPheMetGlnGlyLeuPheTyrrGlyArgAlaGlyGluIleLeu 166  
Db 4109 ATTTCTTTTATACATTTTCTCCTCAGGGTTTCAATTTGGCAAGCTGGAGAGATCCTC 4168  
Qy 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrrGlnAspIleAlaThrPhe 186  
Db 4169 ACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTT 4228  
Qy 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
Db 4229 GATGACCTATAAACACACCTGGAGCATTTGACTACAGGCTCGCCCAATGATCTCTCAA 4288  
Qy 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226  
Db 4289 GTTAAAGGGGCTATAGGTTCCAGGCTTCTGTAATTTACCCAGAAATATAGCAAAATCTGG 4348  
Qy 227 LeuSerValIleIleSerPheIleTyrrGlyTyrrGluMetThrPheLeuIleuSerIle 246  
Db 4349 ACAGAAATAATATATCTCTTCACTATGTTGGCAACTTAACACTGTTACTTTAGCAAT 4408  
Qy 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266  
Db 4409 GTACCCATCATTCATAGCAGGAGTTGTTGAAATGAAATGTTGCTCTGGACAGCACTG 4468  
Qy 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
Db 4469 AAAGATAAGAAAGAACTAGAGGTTCTGGAGAGTCCGCTACTGGAAGCAATAGAAAACTTC 4528  
Qy 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrrGluGluMetLeu 306  
Db 4529 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATATGATGTCAGAGATTG 4588

```

RESULT 7
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Alignment Scores:
Pred. No.: 2,456-218 Length: 9318
Score: 1872.00 Matches: 356
Percent Similarity: 74.22% Conservative: 142
Best Local Similarity: 53.06% Mismatches: 159
Query Match: 56.33% Indels: 14
DB: Gaps: 4

US-09-873-409-1 (1-659) x US-08-793-610-6 (1-9318)
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db 3588 GTCATGTGGAGAAAGAAATCATGTAACATCATGAAGAAGAAGGCAATTACTTCAA 3647
Qy 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34
Db 3648 CTTGTCAACATGCAGACAGCAGCAAAATGAAGTTGAATTAGAAAAATGCAGCTGATGAATCC 3707
Qy 35 GluSerMetThrTyrSerThrGluArgLysThrAnSerLeuProLeuHisSerVal---- 53
Db 3708 AAAAGTGAATTCATGCTTGGAAATGCTTCCTCAAAATGATTCAAGATCCAGTCTTAATAGA 3767

```

```
QY 54 -----lySerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db 3768 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAAGCCCAAGACAGACAGCTTAGTACC 3827
QY 68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db 3828 AAGAGCGCTCGGATGAAGATATACCTCCAGTTTCCTTTGGAGGATTATGAAGCTAAAT 3887
QY 88 LysProGluTrpProPheValValLeuGluGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db 3888 TTAACCTGAATGGCCTTATTTGTGTGTGTATTTGTGTCATTAATAATGGAGCCCTG 3947
QY 108 HisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db 3948 CAACAGCATTTTCAATAATATTTTCAAGATATATAGGGTTTTTCAAGAAATGATGAT 4007
QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db 4008 CCTGAAACAACACGACAGATAGTAACTTGTGTTCATATGTTTCTAGCCCTTGGAAAT 4067
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db 4068 ATTCTTTTATTACATTTTCTTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 4127
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db 4128 ACCAAGCGGCTCCGATACATAGTGTGTTTCCGATCATGCTCAGACAGGATGTGATGTT 4187
QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db 4188 GATGACCTTAAACACCACTGGAGCATTCATCCAGGCTGCCAATGATGCTGCTCAA 4247
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db 4248 GTTAAAGGGCTATAGTTCAGGCTTGCTGTAATATCCAGAAATATAGCAATCTTGG 4307
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db 4308 ACAGGAATAATTATATCTTCATCTATGTTGGCAAGTTTGAACATATGTTACTCTTAGCAAT 4367
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db 4368 GTACCCATCATTCGAATAGCAGGATGTTGAAATGAAATGTTGCTGGACAAGCACTG 4427
QY 267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db 4428 AAGATAAGAAAGAACTAGAGGTGCTGGAGATCGCTACTGACCAATAGAAACTTC 4487
QY 287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db 4488 CGAACCGTTGTTTCTTTGACTCAGGACGAGAAGTTTGAACATATGATGCTCAGAGTTTG 4547
QY 307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db 4548 CAGGTACCATACAGAACTCTTTGAGAAAGACACATCTTTGGAATACATTTTCCTTC 4607
QY 327 SerHisAlaPheIleTyrPheAlaThrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db 4608 ACCAGCGCAATGATGATTTTCTCTATGCTGGATGTTCCGGTTTGGAGCCTACTTGGTG 4667
QY 347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db 4668 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTTCTTGTGT 4727
QY 367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db 4728 GCATGGCCGCGGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAATATCA 4787
QY 387 AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db 4788 GCAGCCCATCATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAA 4847
QY 407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 426
```

```
Db 4848 GGCCTAATGCCGAACACATTTGGAAGGAATGTCAATTTGGTGAAGTTGTATTCAACTAT 4907
QY 427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db 4908 CCCACCCGACCGGACATCCAGTGTCTCAGGACTGAGCTGGAGTGAAGAGGCCGACG 4967
QY 447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValIcInLeuLeuGln 466
Db 4968 ACGCTGGCTCTGCTGGCGCAGTGGCTGTGGAAAGACACAGTGTCTCAGCTCTCTGGAG 5027
QY 467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db 5028 CGGTTCTACGACCCCTTGGCAGGAAGTGTCTGTGTCGCAAGAAATAAAGCGACTG 5087
QY 487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db 5088 AATGTTCTAGTGGCTCGAGCACCTGGGCATCGTGTCCAGGACCCCATCTCTGTTTAC 5147
QY 507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526
Db 5148 TGCAGCATTTGCTGAGAACATTTGCCCTATGGAGACAACACAGCCGGTGTGTACAGGAAG 5207
QY 527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db 5208 ATCGTGAGGGCAGCAAGAGGAGGCCAACATATACATGCTTTCATCGAGTCACTGCTTAATAA 5267
QY 547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db 5268 TATACACTTAAGTAGGAGACANAGGAACCTAGCTCTCTGTGGCCAGAACACAGCAATT 5327
QY 567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db 5328 GCATAGCTCGTGCCCTTGTAGACAGCCTCATATTTTGTCTTGTGATGAAGCCAGTCA 5387
QY 587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db 5388 GCTCTGGATACAGAAAGTGAAGGTTTCCAAAGAGGCTTCCAAAGAGCCCTGGACAAAGCC 5447
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db 5448 CGCACTGCTGATGTGATTTGCTCCCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 5507
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db 5508 GTGTTTCAGATGGCAGATCAAGGAGCATGGCAGCATCAGCAGTGTGTGACAGAAA 5567
QY 647 AspileTyrPheLysLeuValAsnAlaGlnSer 657
Db 5568 GGCATCTATTTTCAATGTCAGTGTCCAGGCT 5600
```

## RESULT 8

```
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechethner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: McDonnell
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,447
FILING DATE: 15-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5994088nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/POCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..424
FEATURE:
NAME/KEY: CDS
LOCATION: 425..4264
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 4265..4669
US-08-752-447-1

```

Alignment Scores:	8.88e-217	Length:	4669
Pred. No.:	1855.00	Matches:	354
Score:	73.47%	Conservative:	139
Percent Similarity:	52.76%	Mismatches:	164
Best Local Similarity:	55.82%	Indels:	14
Query Match:	2	Gaps:	4
DB:			

US-09-873-409-1 (1-659) x US-08-752-447-1 (1-4669)

Qy	1	MetLeuAlaGluIysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	2237	GTCAATTGTGGAGAAAGAAATCATGATGAACATCATGAAGAGAAAGCATTTACTTCAAA	2296
Qy	21	LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet	34
Db	2297	CTTGTCACATGCAGACAGACAGAAATGAAGTTGGAATTAGAAATGCAGCTGATGAATCC	2356
Qy	35	GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal---	53
Db	2357	AAAAGTGAAATTGATGCTTGGGAAATGCTCTTCAAAATGATTCAAGATCCAGTCTAATAAGA	2416
Qy	54	-----LysSerIleLysSerAppPheIleAppLysAlaGluGluSerThr	68
Db	2417	AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACC	2476
Qy	69	---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLysLysLeuAsn	87
Db	2477	AAAGAGGCTCTGAGATGAAGTATACCTCCAGTTTCTTTTGGAGAGATTATGAAGCTAAAT	2536
Qy	88	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	107
Db	2537	TTAACTCAATGGCCCTATTATTTGTGTGTGTATTTTGTGCCATTATAAATGGAGGCCCTG	2596
Qy	108	HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp	126
Db	2597	CAACCCAGCATTTGCATAATATTTTCAAAAGATTATAGGGGTTTTTACAGAATTTGATGAT	2656
Qy	127	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	146
Db	2657	CCTGAACAAACACAGCAGATAGTAACTTGTTTTCATATTGTTCTAGCCCTTGGAAATT	2716
Qy	147	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	166

QY 527 ileLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
Db 3857 ATCGTGAGGCGAGAAAGAGGCGCAACATACATGCTTCATCGAGTCACCTGCTAATAAA 3916  
QY 547 TyrAsnThrClnValGlyLeuLysGlyAlaGlnLeuSerGlyClnLysGlnArgLeu 566  
Db 3917 TATAGCACTAAAGTAGAGAGCAAGAAAGCACTCCTCTGTGTGCCAGAAACAACGCATT 3976  
QY 567 AlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586  
Db 3977 GCATAGCTCGCTGCTGTAGACAGCCTCATATTTTGTGTTGGATGAGCCAGCTCA 4036  
QY 587 AlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
Db 4037 GCTCTGGATACAGAAAGTGAAGAGTTGTCCAGAAAGCCCTGGACAAAGCCAGAGGC 4096  
QY 607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
Db 4097 CGACCTGCAATTGTGATTGCTCACCGCCTGTCACCATCCAGAAATGCAGACTTAATAGTG 4156  
QY 627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646  
Db 4157 GTTTTCAGATGCGAGAGTCAAGAGGATGGCCAGCATCCAGCAGCTGCTGGCACAGAAA 4216  
QY 647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657  
Db 4217 GGCATCTATTTTCAATGGTCACTGTCAGGCT 4249

RESULT 9  
US-09-316-167-1  
; Sequence 1, Application US/09316167  
; Patent No. 6365357  
; GENERAL INFORMATION:  
; APPLICANT: Mechetner, Eugene  
; COMPUTER: Roninson, Igor B  
; TITLE OF INVENTION: Methods and Reagents for Preparing and  
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.  
; STREET: 300 South Wacker Drive, Seventh Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/316,167  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/752,447  
; FILING DATE: 15-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NO. 6365357nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 95,1121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-9808  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4669 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:

NAME/KEY: 5'UTR  
LOCATION: 1..424  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4264  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 4265..4669  
US-09-316-167-1  
Alignment Scores:  
8.88e-217 Length: 4669  
Score: 1855.00 Matches: 354  
Percent Similarity: 73.47% Conservative: 139  
Best Local Similarity: 52.76% Mismatches: 164  
Query Match: 55.82% Indels: 14  
DB: 4 Gaps: 4  
US-09-873-409-1 (1-659) x US-09-316-167-1 (1-4669)  
QY 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
Db 2237 GTCATTTGGAGAAAGAAATCATGATGAATCATGAAGAAAGAAAGGCAATTTACTTCAA 2296  
QY 21 LeuValMetSerGln-----AspIleLysLysAlaAspGluGlnMet 34  
Db 2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCC 2356  
QY 35 GluSerMetTyrSerThrGluArgLysThrAnSerLeuProLeuHisSerVal --- 53  
Db 2357 AAAAGTGAATTTGATGCGCTTGGAAATGCTTCAAAATGATTCAGATCCAGATCTAATAAGA 2416  
QY 54 -----LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68  
Db 2417 AAAAGATCAACTCGTAGAGTGTCGCTGGATCACAAGCCCAACACAGAAAGCTTAGTACC 2476  
QY 69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87  
Db 2477 AAAGAGGCTCGAGTGAAGTATACCTCCAGTTTCTTTGGAGGATTTATGAAGCTAAAT 2536  
QY 88 LysProGluTyrPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107  
Db 2537 TTAACCTGAATGGCCTTATTTTGTGTTGTGTTGTTGTCATTTATTAATGAGGCGCTG 2596  
QY 108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126  
Db 2597 CAACAGCATTTGCAATATATTTTCAAGATATATAGGGTTTTTACAGAAATGATGAT 2656  
QY 127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146  
Db 2657 CCTGAAACCAAAACGACAGATAGTAACCTTGTTTTCACATTGTTTCTAGCCCTTGGAAAT 2716  
QY 147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166  
Db 2717 ATTTCTTTTATACATTTTTCCTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTC 2776  
QY 167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaThrPhe 186  
Db 2777 ACCAAGCGGCTCCGATACATGTTTTCGATCCATGCTACAGGCTCGCCCAATGATGCTCAA 2836  
QY 187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206  
Db 2837 CATGACCCCTAAAAACACACCACCTGGAGCATTTGACTACAGGCTCGCCCAATGATGCTCAA 2896  
QY 207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAnSerMetGly 226  
Db 2897 GTTAAAGGGGCTATAGGTTCCAGGCTTCTGTAATATACCAATATATAGCAAAATCTTGG 2956  
QY 227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246  
Db 2957 ACAGGAATATATATCTTCTATGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016  
QY 247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266

Db 3017 GTACCCATCATTTGCAATAGCAGGAGTTGTGAAATGAAATAATGTTGCTGGCAACGACTG 3076  
Qy |||:||||:||||: |||:||||: |||: |||  
267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286  
|||:||||:||||: |||:||||: |||: |||  
Db 3077 AAGATATAGAAGAACTAGAGGTGCTGGAGATGCTACTGAGCAATAGAAAACCTC 3136  
Qy |||:||||:||||: |||:||||: |||: |||  
287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306  
|||:||||:||||: |||:||||: |||: |||  
Db 3137 CGAACCGTTGTTCTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTG 3196  
Qy |||:||||:||||: |||:||||: |||: |||  
307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326  
|||:||||:||||: |||:||||: |||: |||  
Db 3197 CAGGTACCATACAGAAACTCTTTGAGAAAGCACATCTTTGGAATACATTTTCCTTC 3256  
Qy |||:||||:||||: |||:||||: |||: |||  
327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346  
|||:||||:||||: |||:||||: |||: |||  
Db 3257 ACCCAGGCAATGATGATTTTCCATGCTGGATGTTCCGGTTTGGAGCCCTACTTGGTG 3316  
Qy |||:||||:||||: |||:||||: |||: |||  
347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366  
|||:||||:||||: |||:||||: |||: |||  
Db 3317 GCACATAAATCATGAGCTTTGAGGATGTTCTGTTAGTATTTCAGCTGTTGCTTGGT 3376  
Qy |||:||||:||||: |||:||||: |||: |||  
367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386  
|||:||||:||||: |||:||||: |||: |||  
Db 3377 GCATGGCGGTGGCGAGTCAGTTCAATTTGCTCTGACTATGCCAAAGCCAAATATCA 3436  
Qy |||:||||:||||: |||:||||: |||: |||  
387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406  
|||:||||:||||: |||:||||: |||: |||  
Db 3437 GAGGCCCATCATCATGATCATCTGTAATGAAAACCCCTTTGATTGACAGCTACAGCAGGAA 3496  
Qy |||:||||:||||: |||:||||: |||: |||  
407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426  
|||:||||:||||: |||:||||: |||: |||  
Db 3497 GCCTAATCCGAACACATTTGAGGAAATGTCACATTTTGGTGAAGTTGATTCAACTAT 3556  
Qy |||:||||:||||: |||:||||: |||: |||  
427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446  
|||:||||:||||: |||:||||: |||: |||  
Db 3557 CCCACCCGACCGGACATCCAGTGTCTCAGGACTGAGCTGGAGTGAAGAAGGCCAG 3616  
Qy |||:||||:||||: |||:||||: |||: |||  
447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466  
|||:||||:||||: |||:||||: |||: |||  
Db 3617 ACCTGGCTCTCGTGGGACAGAGTGGCTGGGAAGAGCACAGTGTCCAGCTCTCGGAG 3676  
Qy |||:||||:||||: |||:||||: |||: |||  
467 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 486  
|||:||||:||||: |||:||||: |||: |||  
Db 3677 CGGTTCTACAGCCCTTGCAGGAAAGTGCTGTGATGGCAAGAAATAAGCGACTG 3736  
Qy |||:||||:||||: |||:||||: |||: |||  
487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506  
|||:||||:||||: |||:||||: |||: |||  
Db 3737 AATGTTCACTGGCTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTGTTGAC 3796  
Qy |||:||||:||||: |||:||||: |||: |||  
507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 526  
|||:||||:||||: |||:||||: |||: |||  
Db 3797 TGCAGCATTTGCTGAGAACATTCCTATGAGACAAACAGCCGGGTGTGTCACAGGAAGAG 3856  
Qy |||:||||:||||: |||:||||: |||: |||  
527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546  
|||:||||:||||: |||:||||: |||: |||  
Db 3857 ATCGTGAGGCGAGCAAGAGGCCAACATACATGCTTTCATCGAGTCACTGCTAATAA 3916  
Qy |||:||||:||||: |||:||||: |||: |||  
547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566  
|||:||||:||||: |||:||||: |||: |||  
Db 3917 TATAGCACTAAAGTAGGACAAAGGAACCTCAGCTCTCTGTTGGCCAGAAACAACGCATT 3976  
Qy |||:||||:||||: |||:||||: |||: |||  
567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586  
|||:||||:||||: |||:||||: |||: |||  
Db 3977 GCATAGCTCGTGGCTGTTGACAGCCCTCATATTTGCTTTTGGATGAAGCCAGTCA 4036  
Qy |||:||||:||||: |||:||||: |||: |||  
587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606  
|||:||||:||||: |||:||||: |||: |||  
Db 4037 GCTCTGGATACAGAAAGTGAAGAGTTTCCAAAGAGCCCTCGACAAAGCCAGAGAGGC 4096  
Qy |||:||||:||||: |||:||||: |||: |||  
607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626  
|||:||||:||||: |||:||||: |||: |||

Db 4097 GCACCTGCATTGTGATTGCTCACGGCTGTCACCATCCAGATGCAGACTTAATAGTG 4156  
Qy |||:||||:||||: |||:||||: |||: |||  
627 ValLeuHisAsnGlyLysIleLeuLysGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646  
|||:||||:||||: |||:||||: |||: |||  
Db 4157 GTGTTTCAGATGCGAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAA 4216  
Qy |||:||||:||||: |||:||||: |||: |||  
647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657  
|||:||||:||||: |||:||||: |||: |||  
Db 4217 GGCATCTATTTTCAATGGTCAGTGTCAGGCT 4249  
RESULT 10  
US-09-120-513-1  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Ellens, Harna  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
; TITLE OF INVENTION: SCREENING METHODS THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY:  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,513  
; FILING DATE: 22-JUL-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GP50008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4233 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
US-09-120-513-1  
Alignment Scores:  
Pred. No.: 3,54e-216 Length: 4233  
Score: 1849.50 Matches: 358  
Percent Similarity: 73.13% Conservative: 132  
Best Local Similarity: 53.43% Mismatches: 163  
Query Match: 55.66% Indels: 17  
Gaps: 5  
US-09-873-409-1 (1-659) x US-09-120-513-1 (1-4233)  
Qy 1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20  
|||:||||:||||: |||:||||: |||: |||  
Db 1832 GTCATGTGCGACCAAGGAATCATGAGAGCTCATGAAAGAGAGGCGCATTTTACTTCAAA 1891  
Qy 21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40  
|||:||||:||||: |||:||||: |||: |||



Db 1892 CTTGTCATGACACAG---ACTAGAGGAATGAATTAACACAGGAATTAATGCTTATGAA 1948  
Qy 41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerLeuLysSerAspPhe 60  
Db 1949 TCCCAAAAGTGACACTGGTGCTCTGAGTTTGACTTCAGAAAAATCA---AAATCTCCTTTA 2005  
Qy 61 IleAspLysAla-----GluGluSerThrGln 69  
Db 2006 ATAAGGAGATCAATTGCGAAGATATCCACAGAAAGACACAGACGAGGAGAGAACTTAGT 2065  
Qy 70 SerLysGlu-----IleSerLeuProGluValSerLeuLysLeuLysLeuLysLeu 86  
Db 2066 TCGAAAGAGATGTGATGAAGATGTCCTATGCTTCTTTGGCAGATCCCTAAAGCTA 2125  
Qy 87 AsnLysProGluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 106  
Db 2126 AATATTAGTAGTGAATGGCCCTATTAGTTGGGTGTACTTTGTCTGTATAAATGGGTGC 2185  
Qy 107 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsn 125  
Db 2186 ATACAACCAAGTGTGGCCATAGTGTTCCTCAAGATTTGTAGGGGTTTTTCAAGAGACGAC 2245  
Qy 126 AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 145  
Db 2246 GACATGAACCAACCAACGGAATGTAACTTGTTTCCTCTCTTTCTGGTCATGGGA 2305  
Qy 146 ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyAlaGlyGluIle 165  
Db 2306 ATGATTTCTTTTGTACGTACTCTTTCAAGGCTTCACATTTGCGAAAGCTGGAGAGATC 2365  
Qy 166 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp 185  
Db 2366 CTCACCAAGCGACTCGCATACATGTGCTTCCTCAAAATCCATGCTGCGACAGGATATAAGCTGG 2425  
Qy 186 PheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 205  
Db 2426 TTTGATGACCATAAACACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2485  
Qy 206 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 225  
Db 2486 AATGTTTAAAGGGCTATGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACTT 2545  
Qy 226 GlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuLeuSer 245  
Db 2546 GGCACAGGAATATCTTATCTTGTAGTCTATGGCTGCGCAGCTTACACTTTTACTTGTAGTA 2605  
Qy 246 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 265  
Db 2606 ATTATACCATCTATTGCTGGTGAATTAATTAATGAATCTGTGTGCTGCTGCTGCTGCT 2665  
Qy 266 AsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsn 285  
Db 2666 TTGAAGGACAAAGAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAGAC 2725  
Qy 286 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 305  
Db 2726 TTCGCACTGTTGCTCTTTGCTAGTAGGAGCAGAGATTTGAAACTATGTATGCTGCCAGAGC 2785  
Qy 306 LeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 325  
Db 2786 TTGCAGATCCATCAGAAATGCTTTGAGAAGACACAGCTTTGGGATCATCTTCGCTGCC 2845  
Qy 326 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 345  
Db 2846 TTCACCCAGGCCATGATTATTTTCTATGCTGCTGTTGTTCCGGTTCGGTGCCTACTTG 2905  
Qy 346 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 365  
Db 2906 GTGGCAGCAACTCATGCTTGAAGATGTTAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 2965  
Qy 366 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 385  
Db 2966 GGTGCCATGGCAGCGGAATACAGTTCACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3025

Qy 386 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGln 405  
Db 3026 TCAGCATCCCATCATCATCAGGATCATTCAGAAAATCCCGAGATTGACAGCTACAGCAG 3085  
Qy 406 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 425  
Db 3086 GAGGGCTTGAAGCCCTAATTTGGTTGAAGGAATGTGAATTTAATGGAGTCAATGTTCAAC 3145  
Qy 426 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 445  
Db 3146 TATCCACCCGACCCCAACATCCAGTGCTTCAGGAGCTGAGGTTCCAGGTGAAGAGGGG 3205  
Qy 446 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 465  
Db 3206 CAAACGCTTCGCTGGTGGCAGCAGTGCTCGGGAAGAGTACAGTGTCTCCAGCTGCTC 3265  
Qy 466 GlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGlu 485  
Db 3266 GAGCGCTTCTACAAACCCCATGGCTGGAACAGTGTTCCTAGATGGCAAGAAATAAACAA 3325  
Qy 486 LeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 505  
Db 3326 CTCACGCTCCAGTGGCTCGCGCCACCTGGGCATTTGTCCAGGAGGCCATCTCTGTTT 3385  
Qy 506 AsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAsp 525  
Db 3386 GACTGCGCATCACCGAGAAATCGCTACGGAGACCAACAGCCGCTGCTGCTCATGAG 3445  
Qy 526 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluLysLeuProGlu 545  
Db 3446 GAGATCTGTGAGGCGCCGAGGAGGCCAACATCCACCATTCATCGACTCAGCTGCTGAG 3505  
Qy 546 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyLysGlnArg 565  
Db 3506 AATACACACAGAGTGGGAGACAAAGGACTGAGTGTGCGGGCGGAGAGAGCGCC 3565  
Qy 566 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 585  
Db 3566 ATCGCATCGCGCGCCCTCGTCAGACAGCTCATCTTACTTCTGATGAAGCGACA 3625  
Qy 586 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 605  
Db 3626 TCAGTCTGGATACCGGAGTGAAGGTCTGTCAGGAGCGCTGACAAAGCCAGGGAA 3685  
Qy 606 GlyArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 625  
Db 3686 GGCCGCACTGTGATGTGATCGGCACCGCTGTCCACCATCCAGACCGACACTTGATC 3745  
Qy 626 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 645  
Db 3746 GTGGTGATTCAGAACCGCCAGGTCAAGGAGCAGCGCACCCACAGCAGCTGTGGCCAG 3805  
Qy 646 ArgAspIleTyrPheLysLeuValAsnAla 655  
Db 3806 AAAAGCATCTATTCTTCGATGGTTTCAAGGCT 3835

## RESULT -11

US-09-450-105-1  
; Sequence 1, Application US/09450105  
; Patent No. 6169166  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harna Ellens  
; APPLICANT: John Anthony Feild  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: GP-50008-D1  
; CURRENT APPLICATION NUMBER: US/09/450,105  
; CURRENT FILING DATE: 1999-11-29  
; EARLIER APPLICATION NUMBER: 09/120,513

```

; EARLIER FILING DATE: 1998-07-22
;
; NUMBER OF SEQ ID NOS: 2
;
; SOFTWARE: FastSeq for Windows Version 3.0
;
; SEQ ID NO 1
;
; LENGTH: 4233
;
; TYPE: DNA
;
; ORGANISM: HOMO SAPIENS
;
US-09-450-105-1

```

**Alignment Scores:**

Pred. No.:	3,548-216	Length:	4233
Score:	1849.50	Matches:	358
Percent Similarity:	73.1%	Conservative:	132
Best Local Similarity:	53.4%	Mismatches:	163
Query Match:	55.6%	Indels:	17
DB:	4	Gaps:	5

US-09-873-409-1 (1-659) x US-09-450-105-1 (1-4233)

Qy	1	MetLeuAlaGluIysGlyAlaHislaGluLeuMetAlaLysArgGlyLeuTyrTyrSer	20
Db	1832	GTCAATTTGGCAGCAAGAAATCATGAAGAGCTCATGAAGAAGAAAGGCGATTTACTTCCAA	1891
Qy	21	LeuValMetSerGlnAspIleIysIysAlaaspGluGlnMetGluSerMetThrTyrSer	40
Db	1892	CTTGTGTCATGCACAG---ACTAGAGAAATGAAATTGAACAGGAAATAATGCTTATGAA	1948
Qy	41	ThrGluArgIysThrAsnSerLeuProLeuHisSerValIysSerIleIysSerAspPhe	60
Db	1949	TCCAAAGTGCACCTGGTGCCTGCTGAGTTCCTCGAANAATCA---AAATCTCCTTTA	2005
Qy	61	IleAspIysAla-----GluGluSerThrGln	69
Db	2006	ATAAGGAGATCAATTCGCAGAAGTATCCACAGAACAAGACGAGGAGAAGACTTAGT	2065
Qy	70	SerIysGlu-----IleSerLeuProGluValSerLeuLeuIysIleLeuIysLeu	86
Db	2066	TCGAAAGAGGATGTGGATGAAGATGTGCCTATGTGTTCTCTTTGGCAGATCTCAAAGCTA	2125
Qy	87	AsnIysProGluIleProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr	106
Db	2126	AAATATAGTAAGTGGCCCTATTAGTTGTGGGTGTACTTTGTGCTGTATTAAATGGGTGC	2185
Qy	107	ValHisProValPheSerIleIlePheAlaIysIleIleThrMetPheGly---AsnAsn	125
Db	2186	ATACAACCACTGTTTGGCATAGTGTTCCTCAAGATGTTAGGGGTTTTTCAAGAGACGAC	2245
Qy	126	AspIysThrThrLeuIysHisaspAlaGluIleTyrSerMetIlePheValIleLeuGly	145
Db	2246	GACCATGAACCAACAACAGGAATTGTAACCTTTGTTTTTCCCTTCCTCTTCGTGGTCATGGGA	2305
Qy	146	ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyAtgAlaGlyGluIle	165
Db	2306	ATGATTTCTTTGTAGTACTCTCTTCAAGGCTTCACATTTGGCAAGCTGGAGAGATC	2365
Qy	166	LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp	185
Db	2366	CTCACAAGCGACTCCGATACATGGTCTTCAATCCATCTCGCAGACAGATATAAGCTGG	2425
Qy	186	PheAspGluIysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla	205
Db	2426	TTTGATGACCATAAACAACCACTGCTCGCTGACTCACCGGCTGCTGAGTGCCTTCT	2485
Qy	206	GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet	225
Db	2486	AATGTTAAAGGGGCTATGGGCTCCAGGCTGTGTAGTTTACCAGAAATGTAGCAACCTT	2545
Qy	226	GlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSer	245
Db	2546	GGCAGAGAAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTA	2605
Qy	246	IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla	265



Qy	407	GlyLysPheProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	426
Db	3497	GGCCTAATGCCGAAACACATTCGAAGGAAATGTCATTTGGTGAAGTTGATTCAACTAT	3556
Qy	427	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	446
Db	3557	CCACCCGACCGGACATCCCACTCTTCAGGACATGACCTTGGAGGTGGAAGGCCAG	3616
Qy	447	ThrValAlaPheValGlySerGlyCysGlyLysSerThrSerValGlnLeuGln	466
Db	3617	ACCTGGCTTGGTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCCTGGAG	3676
Qy	467	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	486
Db	3677	CGGTTCTACGACCCCTTGGCAGGGAAGTGGCTGTGATGCCAAGAAATAAAGCCAGCTG	3736
Qy	487	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	506
Db	3737	AAATGTTAGTGGCTCCGACGACCACTTGGGACATCGTGTCCAGGAGCCATCTGTTTGAC	3796
Qy	507	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	526
Db	3797	TGCAGCATTCCTGAGAACATTCCTATGGACACACACCCGGGTGTGTACAGGAAGAG	3856
Qy	527	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	546
Db	3857	ATCGTGGGCGACGAAAGGAGGCCAACATACATGCTTCATCGAGTCACCTGCCTAATAAA	3916
Qy	547	TyrAsnThrClnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	566
Db	3917	TATAGACATTAAGTAGGAGACAAAGAACTCAGCTCTCTGGTGGCCAGAACACGCAAT	3976
Qy	567	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	586
Db	3977	GCCATAGCTCGTCCCTTGTAGACAGCCTCTATTTTGTCTTTGGATGAAGCCACGCTCA	4036
Qy	587	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	606
Db	4037	GCTCTGGATACAGAAAGTGAAAGGTGTGTCAGAAGCCCTGGACAAACCCAGAGAAGGC	4096
Qy	607	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	626
Db	4097	CGCACCTTGCATTTGTGATTTGTCACCGCTGTCCACCATCCAGAATGCAGACTTAATAGTG	4156
Qy	627	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg	646
Db	4157	GTGTTTCAGAAATGGCAGAGTCAAGCAGCATGCGACGCAATCAGCAGCTGCTGGCACAGAA	4216
Qy	647	AspIleTyrPheLysLeuValAsnAlaGln	656
Db	4217	GGCATCTATTTTCAATGGTCAGTGCAG	4246
RESULT 13			
US-08-996-545-1			
; Sequence 1, Application US/08996545			
; Patent No. 5928898			
; GENERAL INFORMATION:			
; APPLICANT: Skatrud, Paul L.			
; APPLICANT: de Waard, Maarten A.			
; APPLICANT: Peery, Robert B.			
; APPLICANT: Andrade, Alan C.			
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of			
; TITLE OF INVENTION: Aspergillus nidulans			
; NUMBER OF SEQUENCES: 3			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Eli Lilly and Company			
; STREET: Lilly Corporate Center			
; CITY: Indianapolis			
; STATE: Indiana			
; COUNTRY: U.S.			
; ZIP: 46285			
; COMPUTER READABLE FORM:			

Db 2434 ATGTTCTTGGTGGTATCATTGATTATACAGCAGTCAACCAATGGTGTGCAATTT 2493  
Qy 160 GlyArgAlaGlyGluLeuThrMetArgLeuArgHisLeuAlaPheYsAlaMetLeu 179  
Db 2494 GCCGTATGCTCCGAGAGACTATTCTGCGCGGAGAAGCACTGCGCTTTCCGAGCATAC 2553  
Qy 180 TyrGlnAspIleAlaTrpPheAspGluYsGluAsnSerThrGlyLeuThrThrIle 199  
Db 2554 CGTCAAGACATTGCTTTCTTGTACAGAGAGAGATAGACCGCGCTCTGACCTCTTC 2613  
Qy 200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219  
Db 2614 CTGTCCACGAGAGAGACATCTCTCGGTGTAGCGGTGTGACTCTAGCAGATCTTG 2673  
Qy 220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleYrGlyTrpGluMet 239  
Db 2674 ATGACCTCCAGCACCTAGGAGCGCTATCATTTATTGCGCTGGCGATTGGTGGAAATTG 2733  
Qy 240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259  
Db 2734 GCCTTAGTTTGTATCTCGGTGTGCGGGTCTCTCGCATGCGGTTCCTACCGATTCTAT 2793  
Qy 260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279  
Db 2794 ATGCTAGCCAGTTTCAATCAGCTCTCAAGCTTGTATTAGGGATCTGCAAACTTTGCT 2853  
Qy 280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299  
Db 2854 TGGCAGGCTACATCGCTATCCGACAGTTTGGCTATTAAACCGGGAAAGGATGCTGG 2913  
Qy 300 GlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319  
Db 2914 GAGATTTACCATGCCAGCTTGACGACAGGCGAGCAGCATCTAATCTCTGCTTTGAGG 2973  
Qy 320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339  
Db 2974 TCATCCTGTATTATGCGTCTCGAGGACACTGTTTCTTCGCGTTCGCGTGGGTTT 3033  
Qy 340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359  
Db 3034 TGTACGGAGGGACACTTCTTGTGTACCCAGCAGTATGACATTTTCGCTTCTTTGTTGT 3093  
Qy 360 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379  
Db 3094 TTCTCCGAGATTCTTCTTGTGTCTCAATCGCGGCGCACCGTCTTTTCTTTCACCCAGAC 3153  
Qy 380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsn 399  
Db 3154 ATGGCAAGCGGAAGATCGCGCGCGCGAATTCGACGAGTTCGACCGAAGCCACAA 3213  
Qy 400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419  
Db 3214 ATTGATACTGTTCTCAAGAGCGGAGAGAGCTCGAACGGTGAAGGTGAATCGAATTT 3273  
Qy 420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439  
Db 3274 AGGAACGTCGACTTCAGATACCCGACCGCCAGAACACAGCTGTCTCGCGCGCTTGGAC 3333  
Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459  
Db 3334 CTGACCGTGAAGCTCGACATAATGTGTGCGTTCGACCCAGCGGTGTGGCAAGAT 3393  
Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479  
Db 3394 ACCACCATGCTGTTCTGAGCGCTTTTACGATGCGATTGCGGGTCCATCTTGTGTAT 3453  
Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499  
Db 3454 GGGAGGACATAGTAACCTAATAATATCACTCTCCTACCGAGCTTTCTGTCACTGGTGCAG 3513  
Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519  
Db 3514 CAGGACCGGACACTGTACCGGGCCACCATCAAGGAACATCTTACTTGTATTGTGCA 3573

Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539  
Db 3574 GATGACGTACCGGAAGAATTTCTTGATTAAAGGCTTCAAGGACGCTAATATCTACGACTTC 3633  
Qy 540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSer 559  
Db 3634 ATCATGTGCTCCCGAGGCGCTTTAATACATGTTGTGGCAGCAAGGAGGCATGTTGCT 3693  
Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579  
Db 3694 GCGGCGCAAAAGCAAGCTGTGCGCATTCGCCGAGCGCTTCTTCGGGATCCCAAAATCCTT 3753  
Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAspSerGluLysValValGlnHisAla 599  
Db 3754 CTTCGTGATGAAGCGACGCTCAGCCCTCAGCTCCGAGTCAGAAAGGTCGTCCAGGCGGCT 3813  
Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619  
Db 3814 TTGGATGCGCTGCCGAGCGCGAACCACCAATCGCGTTGACACCGACTCAGCAGATT 3873  
Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639  
Db 3874 CAAAAGGCGGAGCTTATCTATGTTTTCGACCAAGCAAGATCGTCGAAAGCGGAACGAC 3933  
Qy 640 GlnGluLeuLeuArgAsnAspIleTyrPheLysLeuValAsnAlaGlnSerVal 658  
Db 3934 AGCGAACTGGTCCAGAAAAAGGCGCGGTACTACGAGCTGGTCAACTGCGAGAGCTTG 3990

## RESULT 14

US-08-996-545-3  
; Sequence 3, Application US/08996545  
; Patent No. 5928898

## GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Feery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF INVENTIONS: 3  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/996,545

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Thomas D.

; REGISTRATION NUMBER: 39,872

; REFERENCE/DOCKET NUMBER: X-11766

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3334

; TELEFAX: 317-276-2763

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4002 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: mRNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-996-545-3

## Alignment Scores:

Pred. No.:	4,67e-143	Length:	4002
Score:	1254.00	Matches:	274
Percent Similarity:	58.47%	Conservative:	123
Best Local Similarity:	40.35%	Mismatches:	260
Query Match:	37.74%	Indels:	22
DB:	2	Gaps:	6

US-09-873-409-1 (1-659) x US-08-996-545-3 (1-4002)

Qy 2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21  
Db 1954 AUUGCTGAAACAGGAACUCACCAUGAAUUGUGACCGCGAGGCGCUUACGCCAAACUU 2013  
Qy 22 ValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41  
Db 2014 GUGGAGGCGUACAAGUAUAAGAACAGAAAGGAGGUGAGCCUUGGAGGAGCGCCGACGCU 2073  
Qy 42 GluArgLysThrAsnSer-----LeuProLeuHisSerValLysSerIleLysSerAsp 59  
Db 2074 GAGGAUCUCACGAAGCAGAUUAGCCAAAUCAAAACUGCGUACGCGCAUCACCGAU 2133  
Qy 60 Phe-----IleAspLys-----AlaGluGluSerThrGlnSerLys 71  
Db 2134 CUCGACGGAAAAACCCACACCAUGACCGCACGGGACCCACCAAGUCUGUUCACGCGG 2193  
Qy 72 GluIleSer-----LeuProGluValSerIleLeuLysIleLys 85  
Db 2194 AUUCUUUAAAAAGACCCCGAAACAAACUCCGAAUAUCUAUAUUGGACGCGUCGCAAA 2253  
Qy 86 -----LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 101  
Db 2254 UUGUUGCUUCUUCACACCGCCGUAUAUCCGUAUCUACUACGCGUUCUUCUUA 2313  
Qy 102 ValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleThrMet 121  
Db 2314 GUGUAGCUGGUGGCGGCAACCCACGCAAGCAGUGCUUAUUGCUAAAGCCACGAC 2373  
Qy 122 PheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139  
Db 2374 CUCGCGUCCCGAGAGACAAUAUUGCAAGCUUCGACAUUGCGGAUUCUGGUCUAUG 2433  
Qy 140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159  
Db 2434 AUGUUCUUGUGUUGUAUCAUACUAGUUUAUCAACGACGUAACCAUUGUGUGCAUUU 2493  
Qy 160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179  
Db 2494 GCCGUAGUCUCCGAGAGACUUAUUGCGCGGAGAGACACUGCCUUCGGAGCAUACUC 2553  
Qy 180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 199  
Db 2554 CGUCAAGACAUUGCUUUCUUGACAGGAAGAAUAGCACCGCGCUCUCGACCUUUC 2613  
Qy 200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219  
Db 2614 CUGUACCGCAGCAGCAUGAUCUUCGCGGUGAGCGGUGUAGUCUACGACGAGCAUUG 2673  
Qy 220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 239  
Db 2674 AUGACCCUCCAGCACCGUAGGAGCGGUCAUCAUUAUUGCCUGCGGAUUGGGGAAUUG 2733  
Qy 240 ThrPheLeuLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259  
Db 2734 GCCUAGUUGUAUCUGGUGUGCGGUGUUCUCCGGCAUGCGGUUUCUACCGGAUUCUAU 2793  
Qy 260 AlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAla 279  
Db 2794 AUGUAGCCGAGUUUCAAUCGUCUCCAAAGCUUGUUAUGAGGGAUCUGCAACUUGUCU 2853  
Qy 280 ThrGluAlaLeuGluAlaAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299

Db 2854 UGCAGGCUACAUUGCUUAUCCGCACAGUUGCGUCAUUAACCCGGAAGGGAUGUUCUGG 2913  
Qy 300 GlnMetTyrGluGluMetLeuGlnThrHisArgAsnThrSerLysLysAlaGlnIle 319  
Db 2914 GAGAUUAACCAUGCCAGCUGACGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2973  
Qy 320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339  
Db 2974 UCAUCCUUGUAUUGCGUGCGCAGGACUUGUUAUUGCGUGUUGCGUGCGUGGUGU 3033  
Qy 340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359  
Db 3034 UGUACGAGGAGGACAUUCUUGGUCACACGAGUAUGACAUUUUCCGCUUCUUGUUGU 3093  
Qy 360 PheThrAlaIleAlaTyrGlyValMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379  
Db 3094 UUCUCCGAGAUUCUUCUUGGUGCUCAUCCGCGGCGCAGGAGGAGGAGGAGGAGGAG 3153  
Qy 380 TyrSerLysAlaLysSerGlyAlaHisLeuPheAlaLeuLeuGlyLysLysProAsn 399  
Db 3154 AUGGCAAGGCGAAGAAUGCGCGCCGAAUUCGACGACUGUUGCGAAGGAGGAGGAGG 3213  
Qy 400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419  
Db 3214 AUUGAAUACUGGUCUAGAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3273  
Qy 420 ArgGluValSerPheThrTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439  
Db 3274 AGGAAGUGCUCUACAGAUACCGACCCGCGCAGACAGCAGCAGCAGCAGCAGCAGC 3333  
Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459  
Db 3334 CUGACCGUGAGCGGAGCAUAUUGUUGCGCUUGCGGACCCAGCGGUGUGGCAAGAGU 3393  
Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479  
Db 3394 ACCACAUUGCAUUGCUUGAGCGCUUUAACGAGUUGCGGCGGCAUUCUUGUUGAU 3453  
Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499  
Db 3454 GGAAGGACAUAGUAACUAUAUAUACUCCUACCGCAGCUUCUUGCAGCUGCAGCAGC 3513  
Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519  
Db 3514 CAGGAGCGCACUUGUACCGAGGCGCACCAAGGAGAAACAUUUAUUGGUAUUGCGAA 3573  
Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539  
Db 3574 GAUGACGUACCGGAAGAAUUCUGAUUAAGCUGCAAGGAGCGCUAAUAUACGACUUC 3633  
Qy 540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559  
Db 3634 AUCAUGCGUCCCGGAGGCGGUUAUAUACUUGUUGGAGGAGGAGGAGGAGGAGGAGG 3693  
Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579  
Db 3694 GCGGCGCAAAAGCAACGUGGCGCAUUGCCGAGGCGCUUUCUUGCGGAGUCCCAAAUCCU 3753  
Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599  
Db 3754 CUUCUGAUGAAGCGACGUCACCCUCCAGCAGUAGGAGGAGGAGGAGGAGGAGGAGG 3813  
Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619  
Db 3814 UUGGAGCGCGUCCCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3873  
Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639  
Db 3874 CAAAAGGCGGAGGUUAUUCUUGUUCGACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 3933  
Qy 640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 658

Db 3934 AGCAACUGGUCAGAAAGGCGCGGUACUACGAGGUGGUCACUUGCAGACCUUG 3990

RESULT 15

US-09-328-320-1

; Sequence 1, Application US/09328320

; Patent No. 6228615

; GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.

; APPLICANT: de Waard, Maarten A.

; APPLICANT: Peery, Robert B.

; APPLICANT: Andrade, Alan C.

; TITLE OF INVENTION: Multiple Drug Resistance Gene atpD of

; TITLE OF INVENTION: Aspergillus nidulans

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company

; CITY: Indianapolis

; STATE: Indiana

; COUNTRY: U.S.

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/328,320

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/996,545

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Thomas D.

; REGISTRATION NUMBER: 39,972

; REFERENCE/DOCKET NUMBER: X-11766

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3334

; TELEFAX: 317-276-2763

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4002 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..4002

US-09-328-320-1

Alignment Scores:

Pred. No.: 4,67e-143 Length: 4002

Score: 1254.00 Matches: 274

Percent Similarity: 58.47% Conservative: 123

Best Local Similarity: 40.35% Mismatches: 260

Query Match: 37.74% Indels: 22

DB: 4 Gaps: 6

US-09-873-409-1 (1-659) x US-09-328-320-1 (1-4002)

Qy 2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21

Db 1954 ATTCGTAACAGAACTCACATGTAATGGTTGACCGGAGGCGCTTATCGCAACTT 2013

Qy 22 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41

Db 2014 GTGGAGGCTCAACGATATCAATGAACAGAGGAAGCTGACGCTTGGAGGACCGCAGCT 2073

Qy 42 GluArgLysThrAsnSer-----LeuProLeuHisSerValLysSerIleLysSerAsp 59

Db 2074 GAGGATCTCAGCAATGCGATATTGCCAAATCAAAATCGCTCAAGCGCATCATCCGAT 2133

Qy 60 Phe-----IleAspLys-----AlaGluGluSerThrGlnSerLys 71

Db 2134 CTCGAGGAAACCCACAAACCATTCGCGCAGGCGCACCAAGCTGTGTTCCAGCGG 2193

Qy 72 GluIleSer-----LeuProGluValSerLeuLysLysIleLysLys 85

Db 2194 ATTCCTTTCTAAAGACCCCGGAAACAACTCCGAAATACTCATTATGACGCTGTCTCAA 2253

Qy 86-----LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 101

Db 2254 TTTGTGTCTTCCTTCAACCGCGCTGAAATCCGATCATCTCATCGTCTGTCTCTCTCA 2313

Qy 102 ValLeuAsnGlyThrValHisProValPheSerIleLeuPheAlaLysIleThrMet 121

Db 2314 GGTGTAGCTGGTGGCCCAACCCAGCAAGAGTGTATATGTAAAGCCATCAGCAC 2373

Qy 122 PheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139

Db 2374 CTCTCGCTCCCAAGATCAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTCAITG 2433

Qy 140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159

Db 2434 ATGTTCTTCGTGGTGGTATCATTCAGTTTATACGCGCAGTCAACCAATGGTGTCTCAT 2493

Qy 160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179

Db 2494 GCCGTATGCTCGAGAGACTTATTCGTCGCGGAGAACACTGCGCTTCGGAGATATC 2553

Qy 180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 199

Db 2554 CGTCAAGACATTCCTTTTGACAGAGAGAGATAGACCGCGCTCTGACCTCTTC 2613

Qy 200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219

Db 2614 CTGTCCACCGAGACGAGCATCTCTCCGCTGTAGCGGTGTGACTCTAGGACGATCTTG 2673

Qy 220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 239

Db 2674 ATGACCTCCAGCCCTAGGAGCGCTATCATTTATGCCCTGGCGATGGTGGGAAATG 2733

Qy 240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259

Db 2734 GCCTTAGTTTGTATCTCGGTTGTGCGGTTCTCTCGGATGCGGTTTCTACGATCTAT 2793

Qy 260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279

Db 2794 ATGTAGCCCAAGTTTCAATCAGCTCCAAAGCTTGCTTATGAGGATCTGCAAACTTTG 2853

Qy 280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299

Db 2854 TCGGAGGCTACATCGTCTATCCGACAGTTTCGTCATTAACCCGGGAAAGGATGTCG 2913

Qy 300 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319

Db 2914 GAGATTTACCATGCCAGCTTGACGACCAAGCGAGGACAGCTTAATCTCTGCTTGGAG 2973

Qy 320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339

Db 2974 TCATCCCTGTTTATATGCTGCTCGCAGGCACTTGTCTTCTCGCTGCTCGGTTT 3033

Qy 340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359

Db 3034 TGGTACGAGGAGGACACTTCCTTGGTCACCAAGATGATGACATTTCCCGCTCTTTGTTGT 3093

Qy 360 PheThrAlaIleAlaTyrGlyValaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379

Db 3094 TTCTCCGAGATTCCTTTGGTGTCTCAATCCGCGGCGACCGCTTTTCTTCTTTCACCGAC 3153

Qy 380 TyrSerLysAlaLysSerGlyAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399



Db 3154 ATGGCGAAGCGAAGATGCGCGCGCGCAATTCCGACGACTGTCGACCGGAAGCCACAA 3213  
Qy 400 IleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGluPhe 419  
Db 3214 ATTGATAACTGGTCTGAAGAGCGGAGAGCTCGAAACGGTGGAAAGGTGAATCGAATTT 3273  
Qy 420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439  
Db 3274 AGGAACGTGCACCTTCAGATACCGCCGCGCAGAACACGCTGCTCGCGCGCTTGGAC 3333  
Qy 440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459  
Db 3334 CTGACCGTGAAGCTGGACAATATGTCGGCTTTCGACCCGCGGTGTGGCAAGAGT 3393  
Qy 460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479  
Db 3394 ACCACCATTCGATTGCTTGAGCGCTTTTACGATGCGATTGCGGGTCCATCCTTGTTCAT 3453  
Qy 480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499  
Db 3454 GGGAAAGGACATAAGTAATAACTAAATCACTCTACCGCAGCTTCTGTCTACTGTGTACG 3513  
Qy 500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519  
Db 3514 CAGGAGCCGACACTGTATCCAGGGCCATCAAGGAAACATCTTACTTGGTATTGTTCGAA 3573  
Qy 520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539  
Db 3574 GATGACGTACCGGAAGAATTCTTGATTAAAGCTTCCAAGGACGCTAATATCTACGACTTC 3633  
Qy 540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559  
Db 3634 ATCATGTCTCCCGAGGGCTTTAATACAGTTGTTGGCAGCAAGGAGGCATGTTGTCT 3693  
Qy 560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579  
Db 3694 GCGCGCCAAAAGCAACGTTGTGGCCATTGCCGAGCCCTTCTTCGGGATCCCAAAATCCTT 3753  
Qy 580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599  
Db 3754 CTTCTCGATGAAGCGACGTACGCCCTCGACTCCGAGTCAGAAAAGGTGCTCCAGCGGCT 3813  
Qy 600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619  
Db 3814 TTGGATGCGCTGCCGCGCGCAACCAATCCGCGTTGACACCGACTCAGCACGATT 3873  
Qy 620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639  
Db 3874 CAAAAGCGGACGTTATCTATGTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCAC 3933  
Qy 640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaGlnSerVal 658  
Db 3934 ACGGAATGGTCCAAAAGGCGCGGTACTACGAGCTGGTCACTTGGCAGAGCTTG 3990

Search completed: March 31, 2003, 15:20:38  
Job time : 192.507 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 47.683 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-3

Perfect score: 659

Sequence: 1 MVDENDIRALNVRHYRDHIG.....SESKSAVQALEKDPYRSF 131

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODE=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09873409/runat\_27032003\_115424\_19290/app\_query\_fasta\_1.7544  
-DB=PublishedApplications NA -QWTF=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09873409 @CGN 1.1 960 @runat\_27032003\_115424\_19290  
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:  
10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:  
12: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:  
13: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:  
14: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	1175	10	US-09-873-409-11 Sequence 11, Appl
2	659	100.0	1940	10	US-09-873-409-16 Sequence 16, Appl
3	659	100.0	2021	10	US-09-873-409-15 Sequence 15, Appl
4	619	93.9	2856	10	US-09-873-409-10 Sequence 10, Appl

5	619	93.9	3177	10	US-09-873-409-12 Sequence 12, Appl
6	619	93.9	3621	10	US-09-873-409-14 Sequence 14, Appl
7	619	93.9	3702	10	US-09-873-409-13 Sequence 13, Appl
8	468	71.0	4189	10	US-09-866-866A-3 Sequence 5, Appl
9	467	70.9	4254	10	US-09-917-800A-1424 Sequence 1424, Ap
10	467	70.9	4369	10	US-09-769-097-1 Sequence 1, Appl
11	467	70.9	4425	10	US-09-769-097-3 Sequence 3, Appl
12	465	70.6	4317	9	US-10-044-671-1 Sequence 1, Appl
13	462	70.1	4788	10	US-09-866-866A-7 Sequence 7, Appl
14	461	70.0	3860	10	US-09-866-866A-1 Sequence 1, Appl
15	461	70.0	3860	10	US-09-866-866A-3 Sequence 3, Appl
16	461	70.0	4643	9	US-10-072-621-2 Sequence 2, Appl
17	461	70.0	8630	10	US-09-306-417-1 Sequence 1, Appl
18	461	70.0	8630	10	US-09-306-417-2 Sequence 2, Appl
19	448	68.0	3912	10	US-09-917-800A-1560 Sequence 1560, Ap
20	440	66.8	3924	10	US-09-880-107-2299 Sequence 2299, Ap
21	422	64.0	5010	10	US-09-917-800A-483 Sequence 483, App
22	369	56.0	3512	10	US-09-749-340-7 Sequence 7, Appl
23	368.5	55.9	1076	9	US-09-764-884-13 Sequence 13, Appl
24	368.5	55.9	1076	9	US-10-092-256-13 Sequence 13, Appl
25	368.5	55.9	1367	9	US-09-764-884-21 Sequence 21, Appl
26	368.5	55.9	1367	9	US-10-092-256-21 Sequence 21, Appl
27	366	55.5	4002	9	US-09-758-828-1 Sequence 1, Appl
28	366	55.5	4002	9	US-09-758-828-3 Sequence 3, Appl
29	364	55.2	1810	10	US-09-749-340-4 Sequence 4, Appl
30	364	55.2	4175	10	US-09-749-340-3 Sequence 3, Appl
31	363	55.1	3861	9	US-09-938-842A-263 Sequence 263, App
32	357	54.2	4653	9	US-10-101-388-2 Sequence 2, Appl
33	356	54.0	981	10	US-09-770-445-272 Sequence 272, App
34	350	53.1	216	10	US-09-864-761-27861 Sequence 27861, A
35	350	53.1	465	10	US-09-864-761-11237 Sequence 11237, A
36	347	52.7	939	10	US-09-974-300-6511 Sequence 6511, Ap
37	339	51.4	2681	10	US-09-749-340-8 Sequence 8, Appl
38	335	50.8	3861	9	US-09-938-842A-2237 Sequence 2237, Ap
39	333	50.5	2698	10	US-09-749-340-5 Sequence 5, Appl
40	327	49.6	53165	9	US-09-939-964-1 Sequence 1, Appl
41	326	49.5	8036	9	US-10-101-388-1 Sequence 1, Appl
42	325	49.3	2066	10	US-09-873-409-9 Sequence 9, Appl
43	319	48.4	2298	9	US-10-156-239-18 Sequence 18, Appl
44	319	48.4	2298	10	US-09-795-693-18 Sequence 18, Appl
45	319	48.4	3408	9	US-10-156-239-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-09-873-409-11  
; Sequence 11, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 1175  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-11

Alignment Scores:  
Pred. No.: 1.58e-81 Length: 1175  
Score: 659.00 Matches: 131  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-11 (1-1175)

Qy	1	MetValAspGluIleAspIleAArgAlaLeuAsnValAArgHisTyrAArgAspHisIleGly	20
Db	418	ATGTTGGATGAGAATGACATCAGAGCTTTAAATGTGGCATTATCGAGACCATATGGA	477
Qy	21	ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleIleVsTyrGly	40
Db	478	GTGTTTAGTCAAGAGCCTGTTTGTTCGGACCAACCATCAGTAACCAATATCAAGTATGGA	537
Qy	41	ArgAspAspValThrAspGluGluMetGluArgAlaAalaArgGluAlaAsnAlaTyrAsp	60
Db	538	CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAACCAATTCGTATGAT	597
Qy	61	PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMet	80
Db	598	TTTATCATCGAGTTTCCTAATAAATTTAATACATTGTTAGCGGAAAAAGCAGCTCAAAATG	657
Qy	81	SerGlyGlyGlnLysGlnAArgIleAlaIleAlaAargAlaLeuValAArgAsnProLysIle	100
Db	658	AGTCAGCGGCAGAACACAGAGGATCGCAATTTGCTCGTCCTTAGTTTCGAACCCCAAGATT	717
Qy	101	LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla	120
Db	718	CTGATTTTAGATGAGGCTACGTCCTGCCCTGATTCCAGAAAGCAAGTCAGCTGTTCAAGCT	777
Qy	121	AlaLeuGlnLysAspThrProAArgTyrSerPhe	131
Db	778	GCATCGAGAGGATACCCCGAGGTATTCATTT	810

## RESULT 2

```

US-09-873-409-16
; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 16
; LENGTH: 1340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-16

```

Alignment Scores:	3.21e-81	Length:	1940
Pred. No.:	Score:	Matches:	131
	659.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	10		

US-09-873-409-3 (1-131) x US-09-873-409-16 (1-1940)

Qy	1	MetValAspGluAsnAspLeuArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20
Db	1183	ATGTTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA	1242
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleIysTyrGly	40
Db	1243	GTGGTTAGTCAAGAGCGCTGTTTGTTCGGAGCCACCATCAGTAACAATATCAAGTATGGA	1302
Qy	41	ArgAspAspValThrAspGluGluMetGluArgAlaAaTgGluAaAsnAlaTyrAsp	60
Db	1303	CGAGATGATGTGACTGATGAAGAGATCGGAGAGAGCAGCAAGGGAAGCAAAATCGCTATGAT	1362
Qy	61	PheIleMetGluPheProAsnIysPheAsnThrLeuValGluIysGluValAlaGlnMet	80

Db	1363	TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTACGGGAAAAGGAGCTCAAAATG	1422
Qy	81	SerGlyGlyGlnGlyGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProGlySile	100
Db	1423	AGTGGAGGGCAGAAAACAGAGGATCCCAATTCTCGTGCCTTAGTTTCGAAACCCCAAGATT	1482
Qy	101	LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAla	120
Db	1483	CTGATTTTAGATGAGGCTACGCTCGCCCTGGATTTCAGAAACCAAGTCAGCTGTTCAAGCT	1542
Qy	121	AlaLeuGluIysAspThrProArgTyrSerPhe	131
Db	1543	GCACCTGGAGAAGGATACCCCAAGGTATTCATTT	1575

### RESULT 3

```

US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 2021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

```

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:

US-09-873-409-3 (1-131) x US-09-873-409-15 (1-2021)

Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20
Db	1264	ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGCATATTATCAGAGCCATATTGGA	1323
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly	40
Db	1324	GTGGTTAGTCAAGAGCGCTTTTGTTCGGGACCAACCATCAGTAAACAATATCATCAATATGGA	1383
Qy	41	ArgAspAspValThrAspGluGluMetGluArgAlaIaIaArgGluAlaAsnAlaTyrAsp	60
Db	1384	CGAGATGATGTGACTGATCAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAATGCCGTATGAT	1443
Qy	61	PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValaIaGlnMet	80
Db	1444	TTTATCATGGAGTTTCTTAATTAATTTAATCATTTGGTATGGGGGAANAAGAGCTCAATG	1503
Qy	81	SerGlyGlyGlnLysGlnArgIleAlaIaIaArgAlaLeuValArgAsnProLysIle	100
Db	1504	AGTGGAGGCGAGAAACAGAGGATCGCAATTGCTGTGCTTAGTTTCGAAACCCCAAGATT	1563
Qy	101	LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla	120
Db	1564	CTGATTTTATATCAGAGCTACGCTCGCCCTGGATTCAGAAAGCAAGTACAGCTGTTCAGACT	1623

QY 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131  
|||||  
Db 1624 GCACGGAGAGGATACCCCGAGGTATTCATT 1656

## RESULT 4

US-09-873-409-10

; Sequence 10, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 10

; LENGTH: 2856

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-873-409-10

Alignment Scores:

Pred. No.:	1,84e-75	Length:	2856
Score:	619.00	Matches:	124
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.93%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-3 (1-131) x US-09-873-409-10 (1-2856)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20

Db 418 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATTTGA 477

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40

Db 478 GTGGTTAGTCAGAGCTGTTTGTTCGGGACCAACCATCATGTAACATATCAAGTATGA 537

QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60

Db 538 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 597

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMet 80

Db 598 TTTATCATGAGGTTTCTTAATAATTTAATACATTTGTTAGGGGAAAAGGAGCTCAAATG 657

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100

Db 658 AGTGGAGGGCAGAAACAGAGGATCGCAATTCGTGCGCTTAGTTCCAAACCCCAAGATT 717

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120

Db 718 CTGATTTTAGATGAGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

QY 121 AlaLeuGluLys 124

Db 778 GCACCTGGAGAAG 789

RESULT 5

US-09-873-409-12

; Sequence 12, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 12

; LENGTH: 3177

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: No. US20020037522A1e

; LOCATION: (198)..(198)

; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)

US-09-873-409-12

Alignment Scores:

Pred. No.:	2,14e-75	Length:	3177
Score:	619.00	Matches:	124
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.93%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-3 (1-131) x US-09-873-409-12 (1-3177)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20

Db 739 ATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATTTGA 798

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40

Db 799 GTGGTTAGTCAGAGCTGTTTGTTCGGGACCAACCATCATGTAACATATCAAGTATGA 858

QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60

Db 859 CGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 918

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMet 80

Db 919 TTTATCATGAGGTTTCTTAATAATTTAATACATTTGTTAGGGGAAAAGGAGCTCAAATG 978

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100

Db 979 AGTGGAGGGCAGAAACAGAGGATCGCAATTCGTGCGCTTAGTTCCAAACCCCAAGATT 1038

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120

Db 1039 CTGATTTTAGATGAGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1098

QY 121 AlaLeuGluLys 124

Db 1099 GCACCTGGAGAAG 1110

RESULT 6

US-09-873-409-14

; Sequence 14, Application US/09873409

; Patent No. US20020037522A1

; GENERAL INFORMATION:

; APPLICANT: Frank, Markus

; APPLICANT: Sayegh, Mohamed

; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein

; FILE REFERENCE: 81994/268611

; CURRENT APPLICATION NUMBER: US/09/873,409

; CURRENT FILING DATE: 2001-06-05

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 14

; LENGTH: 3621

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-873-409-14

Alignment Scores:

Pred. No.:	2,57e-75	Length:	3621
------------	----------	---------	------

Score: 619.00 Matches: 124  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 93.93% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-14 (1-3621)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
DB 1183 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGCGCATTTATCGAGACCATATTGGA 1242  
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
DB 1243 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCCACCATCAGTAACAATAATCAAGTATGGA 1302  
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
DB 1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 1362  
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
DB 1363 TTTATCATGGAGTTTCTTAATAATTTAATACATTTGTTAGGGGAAAAGAGCTCAAAATG 1422  
QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
DB 1423 AGTGGAGGCGAGAAACAGAGATCGCAATTTGCTCGTCCCTTAGTTCGAAACCCCAAGATT 1482  
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
DB 1483 CTGATTTTAGATGAGGCTAGCTGCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542  
QY 121 AlaLeuGluLys 124  
DB 1543 GCACGTGAGAAG 1554

## RESULT 7

US-09-873-409-13  
; Sequence 13, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 3702  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (723) - (723)  
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)  
US-09-873-409-13

## Alignment Scores:

Pred. No.: 2,65e-75 Length: 3702  
Score: 619.00 Matches: 124  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 93.93% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-873-409-13 (1-3702)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
DB 1264 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGCGCATTTATCGAGACCATATTGGA 1323

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
DB 1324 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCCACCATCAGTAACAATAATCAAGTATGGA 1383  
QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
DB 1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 1443  
QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
DB 1444 TTTATCATGGAGTTTCTTAATAATTTAATACATTTGTTAGGGGAAAAGAGCTCAAAATG 1503  
QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
DB 1504 AGTGGAGGCGAGAAACAGAGATCGCAATTTGCTCGTCCCTTAGTTCGAAACCCCAAGATT 1563  
QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
DB 1564 CTGATTTTAGATGAGGCTAGCTGCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623  
QY 121 AlaLeuGluLys 124  
DB 1624 GCACGTGAGAAG 1635

## RESULT 8

US-09-866-866A-5  
; Sequence 5, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 4189  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-866-866A-5

## Alignment Scores:

Pred. No.: 2.2e-54 Length: 4189  
Score: 468.00 Matches: 90  
Percent Similarity: 91.06% Conservative: 22  
Best Local Similarity: 73.17% Mismatches: 11  
Query Match: 71.02% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-866-866A-5 (1-4189)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
DB 1354 ATCGACGGACAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCAATTTGGTGTG 1413  
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
DB 1414 GTGAGTCAGGAACCTGTGCTGTTTGCACCATCGATCGCGAGAACCATTCGCTATGCCCGA 1473  
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
DB 1474 GAAGATGTCCCATCATGATGAGATTGAGAAGCTGTCAAGGAAGCCCAATGCTATGACTTC 1533  
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Db 1534 ATCATGAACTGCCCAATTTGACCCCTGGTTGGTGAGAGAGGGCGCAGCTGAGT 1593  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsnProLysIleLeu 101  
Db 1594 GGGGGACAGAAACAGAGATCGCATTTGCCGGGCGCTGGTCCGCATCCCAAGATCCTT 1653  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
Db 1654 TTGTTGGAGCGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGGTGTGAGGCGCGA 1713  
Qy 122 LeuGluLys 124  
Db 1714 CTGGATAAG 1722  
RESULT 9  
US-09-917-800A-1424  
; Sequence 1424, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1424  
; LENGTH: 4254  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855  
US-09-917-800A-1424  
Alignment Scores:  
Pred. No.: 3.09e-54 Length: 4254  
Score: 467.00 Matches: 90  
Percent Similarity: 91.06% Conservative: 22  
Best Local Similarity: 73.17% Mismatches: 11  
Query Match: 70.86% Indels: 0  
DB: 10 Gaps: 0  
US-09-873-409-3 (1-131) x US-09-917-800A-1424 (1-4254)  
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
Db 1456 ATTCAGCGACAGACATCAGACCATCAATGTGAGGTATCTCGGGAAATCATTTGGGGTG 1515  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41

Db 1516 GTGAGTCAGAAACCCGTGCTGTTGCCACCAAGTTCGCGAAACATTCGCTATGCGCGA 1575  
Qy 42 AspAspValThrAspGluMetGluArgAlaAArgGluAlaAsnAlaTyrAspPhe 61  
Db 1576 GAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAGCAATGCTTATGACTTC 1635  
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
Db 1636 ATCATGAACTGCCCAATTTGACACCTGCTGGTTGAGAGAGGGCGCAGCTGAGT 1695  
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsnProLysIleLeu 101  
Db 1696 GGGGGACAGAAACAGAGATCGCCATTGCCGGGCGCTGGTCCGCAACCCCAAGATCCTT 1755  
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
Db 1756 TTGTTGGATGAGCCACGTCAGCCTTGGACACAGAAAGCGAGCCGTGCTTACAGGCGCT 1815  
Qy 122 LeuGluLys 124  
Db 1816 CTGGATAAG 1824  
RESULT 10  
US-09-769-097-1  
; Sequence 1, Application US/09769097  
; Patent No. US20020055128A1  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Ellens  
; APPLICANT: John Anthony Feild  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: GP-50009-C2  
; CURRENT APPLICATION NUMBER: US/09/769,097  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 09/208,809  
; PRIOR FILING DATE: 1998-12-09  
; PRIOR APPLICATION NUMBER: 09/156,800  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: US99/20770  
; PRIOR FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 4369  
; TYPE: DNA  
; ORGANISM: RATTUS RATTUS  
US-09-769-097-1  
Alignment Scores:  
Pred. No.: 3.21e-54 Length: 4369  
Score: 467.00 Matches: 90  
Percent Similarity: 91.06% Conservative: 22  
Best Local Similarity: 73.17% Mismatches: 11  
Query Match: 70.86% Indels: 0  
DB: 10 Gaps: 0  
US-09-873-409-3 (1-131) x US-09-769-097-1 (1-4369)  
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
Db 1684 ATTCAGCGACAGACATCAGACCATCAATGTGAGGTATCTCGGGAAATCATTTGGGGTG 1743  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
Db 1744 GTGAGTCAGAAACCCGTGCTGTTGCCACCAATTCGCGAAACATTCGCTATGCGCGA 1803  
Qy 42 AspAspValThrAspGluMetGluArgAlaAArgGluAlaAsnAlaTyrAspPhe 61  
Db 1804 GAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAGCAATGCTTATGATTTC 1863





```
QY 122 LeuGlulys 124
Db 1789 CTGGATAAG 1797

RESULT 13
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:
Pred. No.: 1.79e-53 Length: 4788
Score: 462.00 Matches: 89
Percent Similarity: 90.24% Conservative: 22
Best Local Similarity: 72.36% Mismatches: 12
Query Match: 70.11% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-866-866A-7 (1-4788)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1345 ATCGAGGACAGACATCAGAACCATCATGTGAGGTATCTGAGGAGATCATTTGGTGTG 1404
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1405 GTGAGTCAGGAACCTGTGCTGTTTGCACCAACCATCGCCGAGAACATTCGCTATGCCGA 1464
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1465 GAAGATGTCCACCATGATGAGATGTGAGAAAGCTGTCAAGGAGCCCAATGCTATGACTTC 1524
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1525 ATCATGAAGCTGCCCAACCAATTGTACACCTGTTGGTGAGAGAGGGCGACGTGAGT 1584
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1585 GGGGGACAGAAACAGAGATCGCCATTCGCCGGCCCTGTCGCCAATCCAGATCTT 1644
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAla 121
Db 1645 TTGTTGGAGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTTTCAGGCCGA 1704
QY 122 LeuGlulys 124
Db 1705 CTGGATAAG 1713

RESULT 14
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.: 1.82e-53 Length: 3860
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-866-866A-1 (1-3860)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1357 GTTGATGGACAGATATTAGGACCATTAATGTAAGGTTTCTACGGGAAATCATTTGGTGTG 1416
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1417 GTGAGTCAGGAACCTGTGTTTGTCCACCAACATAGCTGAAACACATTCGCTATGCCGT 1476
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1477 GAAATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAGCCCAATGCTATGACTTT 1536
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1537 ATCATGAAGCTGCCCAACCAATTGTACACCTGTTGGTGAGAGAGGGCGCCAGTTGAGT 1596
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1597 GGTGGCAGACAGACAGGATCGCCATTCGACCTGCTGCTGTCGCAACCCCAAGATCTTC 1656
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAla 121
Db 1657 CTGCTGATGAGCCACCTCAGCCCTTGGACACAGAAAGCGAAGCAGTGTTCAGGTGCT 1716
QY 122 LeuGlulys 124
Db 1717 CTGGATAAG 1725

RESULT 15
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-3
```

; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 3860  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-866-866A-3

Alignment Scores:  
Pred. No.: 1 82e-53 Length: 3860  
Score: 461.00 Matches: 90  
Percent Similarity: 89.43% Conservative: 20  
Best Local Similarity: 73.17% Mismatches: 13  
Query Match: 69.95% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-3 (1-131) x US-09-866-866A-3 (1-3860)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
DB 1357 GTTCATGCAGAGATATTAGACCATTAATGTAAAGTTTCTACGGGAATCATTTGGTGTG 1416  
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
DB 1417 GTGAGTCAGGAACCTGTATTGTTTGCACCACCATAGCTGAAAAACATTCGCTATGGCGGT 1476  
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
DB 1477 GAAATGTCACCATGTGATGAGATTGAGAAAGCTGTCAAGGAAGCCATGCTATGACTTT 1536  
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
DB 1537 ATCATGAACCTGCTCATTAATTTGACACCTGTTGGAGAGAGAGGGCCCGAGTTGAGT 1596  
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
DB 1597 GGTGGGCAGAAAGCAGAGGATCGCCATTGCACCTGCTGGTTCCCAACCCCAAGATCCTC 1656  
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
DB 1657 CTGCTGATGAGGCCACGTGAGCCTTGGACACAGAAAGCAGAGTGTTTCAGGTGGCT 1716  
QY 122 LeuGluLys 124  
DB 1717 CTGATTAAG 1725

Search completed: March 31, 2003, 16:00:29  
Job time : 58.683 secs

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 19.1843 Seconds  
(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-3

Perfect score: 659

Sequence: 1 MVENDIRALNVRHYRDHIG.....SEKSAVAQALEKQTPRYSF 131

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh

-Q/cgn2\_1/USFTO\_spool/US09873409/runat\_27032003.115422.19259/app\_query.fasta\_1.7544  
-DB=Issued Patents NA -QFT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CNC 1.1 578 @runat\_27032003.115422.19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCUTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	71.6	4233	3	US-09-120-513-1
2	472	71.6	4233	3	US-09-450-105-1
3	461	70.0	4264	2	US-08-784-649A-1
4	461	70.0	4264	2	US-08-784-649A-5
5	461	70.0	4646	1	US-08-181-471-2
6	461	70.0	4669	2	US-08-752-447-1
7	461	70.0	4669	4	US-09-316-167-1
8	461	70.0	4669	6	5206352-3
9	461	70.0	6505	2	US-08-793-610-5
10	461	70.0	9318	2	US-08-793-610-6
11	449	68.1	4669	2	US-08-583-276-18
12	372	56.4	2726	1	US-08-461-823-1

13	366	55.5	4002	2	US-08-996-545-1	Sequence 1, Appli
14	366	55.5	4002	2	US-08-996-545-3	Sequence 3, Appli
15	366	55.5	4002	4	US-09-328-320-1	Sequence 1, Appli
16	366	55.5	4002	4	US-09-328-320-3	Sequence 3, Appli
17	360	54.6	3924	1	US-08-395-246C-1	Sequence 1, Appli
18	360	54.6	4047	2	US-08-612-734B-1	Sequence 1, Appli
19	355	53.9	4224	1	US-08-612-521-1	Sequence 1, Appli
20	346.5	52.6	4800	2	US-08-612-734B-3	Sequence 3, Appli
21	325.5	49.4	2376	1	US-08-394-880B-1	Sequence 1, Appli
22	319	48.4	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
23	309	46.9	3924	2	US-08-996-644-3	Sequence 3, Appli
24	309	46.9	3924	3	US-09-352-552-3	Sequence 3, Appli
25	309	46.9	3927	2	US-08-996-644-1	Sequence 1, Appli
26	309	46.9	3927	3	US-09-352-552-1	Sequence 1, Appli
27	300.5	45.6	6143	1	US-08-612-521-3	Sequence 3, Appli
28	276	41.9	2061	4	US-09-061-764A-17	Sequence 17, Appl
29	275	41.7	6492	4	US-08-961-527-188	Sequence 188, App
30	272	41.3	3576	4	US-08-976-259-79	Sequence 79, Appl
31	271	41.1	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
32	270	41.0	775	4	US-09-221-017B-860	Sequence 860, App
33	270	41.0	1959	3	US-09-061-764A-4	Sequence 4, Appli
34	268	40.7	7721	3	US-08-772-270A-14	Sequence 14, Appl
35	267.5	40.6	3792	4	US-09-351-224E-10	Sequence 10, Appl
36	266	40.4	5120	3	US-08-772-270A-6	Sequence 6, Appli
37	266	40.4	8370	2	US-08-488-706-1	Sequence 1, Appli
38	265	40.2	7760	4	US-08-961-527-63	Sequence 63, Appli
39	265	40.2	28804	2	US-08-592-874-1	Sequence 1, Appli
40	265	40.2	28804	3	US-09-096-942-2	Sequence 2, Appli
41	265	40.2	28804	3	US-09-096-867-2	Sequence 2, Appli
42	265	40.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
43	265	40.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
44	264.5	40.1	3909	1	US-08-232-537-1	Sequence 1, Appli
45	262.5	39.8	3999	4	US-09-351-224E-9	Sequence 9, Appli

#### ALIGNMENTS

RESULT 1  
US-09-120-513-1  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Ellens, Harma  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
; TITLE OF INVENTION: SCREENING METHODS THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY:  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,513  
; FILING DATE: 22-JUL-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GPS0008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015

```

; TELEFAX: 610-270-5090
;
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 1
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 433 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: Genomic DNA
;
; US-09-120-513-1

```

Alignment Scores:	
Pred. NO.:	3.59e-55
Score:	472.00
Percent Similarity:	91.06%
Best Local Similarity:	73.98%
Query Match:	71.62%
DB:	3
Length:	4233
Matches:	91
Conservative:	21
Mismatches:	11
Indels:	0
Gaps:	0

US-09-873-409-3 (1-131) x US-09-120-513-1 (1-4233)

Qy	2	ValAspGluAsnAspIleAlaArgAlaLeuAsnValAlaRgHisTyrArgAspHisIleGlyVal	21
Db	1376	ATCGACGGCAGGACATCAGGCATCAATGTGAGGATTCTCGCGGAATCATTTGGGGTG	1435
Qy	22	ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrglyArg	41
Db	1436	GTGAGTCAGAACCCGTGCTTTCGCCACGATGCGGAAAACATTTCGTATGGCCGA	1495
Qy	42	AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe	61
Db	1496	GAAACGTCACCATGGATGAGTAGAAGAAGCTGCAAGGAAGCCAATGCCATGACTTC	1555
Qy	62	IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer	81
Db	1556	ATCATGAACATGCCCCACAAATTTAAACCCCTGGTTGTGTGAGAGGGGGCGCAGCTAGT	1615
Qy	82	GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu	101
Db	1616	GGGGGACAGAAACAGAGGATCGCCATTGCCGGGGCCCTGGTCCGCAACCCCAAGATCCCT	1675
Qy	102	IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla	121
Db	1676	TTCGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAGCCGTTCAGGCCGCT	1735
Qy	122	LeuGluLys	124
Db	1736	CTGGATAAG	1744

## RESULT 2

```

US-09-450-105-1
/ Sequence 1, Application US/09450105
/ Patent No. 6169166
/ GENERAL INFORMATION:
/ APPLICANT: Kimberly Anne Brun
/ APPLICANT: Richard James Cheney
/ APPLICANT: Harma Ellens
/ APPLICANT: John Anthony Feild
/ APPLICANT: Lin Yue
/ TITLE OF INVENTION: POLYNUCLEOTIDE
/ TITLE OF INVENTION: ENCODING RAT M
/ FILE REFERENCE: GP-50008-D1
/ CURRENT APPLICATION NUMBER: US/09/
/ CURRENT FILING DATE: 1999-11-29
/ EARLIER APPLICATION NUMBER: 09/120
/ EARLIER FILING DATE: 1998-07-22
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: FastSeq for Windows Vers
/ SEQ ID NO 1
/ LENGTH: 4233
/ TYPE: DNA
/ ORGANISM: HOMO SAPIENS
US-09-450-105-1

```

Alignment Scores:

Pred. No.:	3,596-55	Length:	4233
Score:	472.00	Matches:	91
Percent Similarity:	91.06%	Conservative:	21
Best Local Similarity:	73.98%	Mismatches:	11
Query Match:	71.62%	Indels:	0
DB:	4	Gaps:	0

US-09-873-409-3 (1-131) x US-09-450-105-1 (1-4233)

Qy	2	ValAspGluAsnAspIleA	gAlaLeuAsnValA	ArgHisTyrA	rgAspHisIleG	Val 21
Db	1376	ATCGACGGACAGGACAT	CAGGACCAATGTGAG	GAGTATCTCGCGGA	AAATCATTTGGGGTG	1435
Qy	22	ValSerGluProValIleu	PheGlyThrThrIle	SerAsnAsnIle	LyTyrGlyA	rg 41
Db	1436	GTGAGTCAGGAACCGG	TGTTTCCACCACGAT	GTCCGAAAAACATTC	GCCTATGGCCGA	1495
Qy	42	AspAspValThrAspGlu	MetGluA	rgAlaA	ArgGluA	AlaAsnAlaTyrAspPhe 61
Db	1496	GAAAACTACCATCGAT	GAGATAGAAAGCTG	CACAGGAAGCCATGC	CTATGACTTC	1555
Qy	62	IleMetGluPheProAsn	lysPheAsnThrIleu	ValGlyGluLysGly	A	AlaGlnMetSer 81
Db	1556	ATCATGAACATGCCCC	CACAAATTTAAACCC	CTGGTTGGTGAGAGGG	CGCGCAGCTCAGT	1615
Qy	82	GlyGlyGlnLysGlnA	rgIleAlaIleA	rgAlaLeuValA	rgAsnProLysIleLeu 101	
Db	1616	GGGGGACAGAAACAG	AGGATGCCATTC	CCCGGGGCCCTGGTCC	GACACCCCAAGATCCCT	1675
Qy	102	IleLeuAspGluAlaThr	SerAlaLeuAspSer	GluSerLysSerAla	ValGlnAlaA	la 121
Db	1676	TTGTTGGATGGGCCAC	CGTCAGCCTTGGAC	A	CAGAAACGGAAGCCG	TGTCAGCCGCT 1735
Qy	122	LeuGluLys	124			
Db	1736	CTGGATAAG	1744			

### RESULT 3

```

US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg. No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:

```



```
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/181,471
/ FILING DATE: 13-JAN-1994
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/041,553
/ FILING DATE: 02-APR-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fitting, Thomas
/ REGISTRATION NUMBER: 34,163
/ REFERENCE/DOCKET NUMBER: ANT0029P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619-792-3680
/ TELEFAX: 619-792-8477
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4646 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 425..4267
/ US-08-181-471-2

Alignment Scores:
Pred. No.: 1,37e-53 Length: 4646
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 1 Gaps: 0

US-09-873-409-3 (1-131) x US-08-181-471-2 (1-4646)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1781 GTTGATGCAGCAGGATATTAGGACCAATAATGTAAGTTTCTACGGGAATCAATGGTG 1840
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1841 GTGAGTCAGGAACCTGTATTGTTTGCACCACGATAGCTGAAACATTCGCTATGGCCGT 1900
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1901 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAGGCCAATGCCTATGACTTT 1960
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1961 ATCATGAACTGCCTCATAAATTTGACACCCCTGTTGAGAGAGAGGGGCCAGTTGAGT 2020
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 2021 GGTGGGAGAGAGAGAGATCGCCATTCAGTCCTGTTGTCACACCCCAAGATCCCTC 2080
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 2081 CTGCTGGATGAGGCCACGTCAGCCTTGACACAGAAAGCGAGCAGTGGTTTCAGGTGCT 2140
QY 122 LeuGluLys 124
DB 2141 CTGGATAAG 2149

RESULT 6
US-08-752-447-1
/ Sequence 1, Application US/08752447
/ Patent No. 5994088
/ GENERAL INFORMATION:

/ APPLICANT: Mechetner, Eugene
/ APPLICANT: Roninson, Igor B
/ TITLE OF INVENTION: Methods and Reagents for Preparing and
/ TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
/ STREET: 300 South Wacker Drive, Seventh Floor
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/752,447
/ FILING DATE: 15-NOV-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 5994088nan, Kevin E
/ REGISTRATION NUMBER: 35,303
/ REFERENCE/DOCKET NUMBER: 95,1121
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-913-0001
/ TELEFAX: 312-913-9808
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4669 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: 5' UTR
/ LOCATION: 1..424
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 425..4264
/ FEATURE:
/ NAME/KEY: 3' UTR
/ LOCATION: 4265..4669
/ US-08-752-447-1

Alignment Scores:
Pred. No.: 1,38e-53 Length: 4669
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 2 Gaps: 0

US-09-873-409-3 (1-131) x US-08-752-447-1 (1-4669)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1781 GTTGATGCAGCAGGATATTAGGACCAATAATGTAAGTTTCTACGGGAATCAATGGTG 1840
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1841 GTGAGTCAGGAACCTGTATTGTTTGCACCACGATAGCTGAAACATTCGCTATGGCCGT 1900
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1901 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAGGCCAATGCCTATGACTTT 1960
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1961 ATCATGAACTGCCTCATAAATTTGACACCCCTGTTGAGAGAGAGGGGCCAGTTGAGT 2020
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
```



```
Db 2021 GGTGGGCAAGCAGAGGATCGCATTCGACGCGCCCTGTTGCGAACCCCAAGATCCTC 2080
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAlaA 121
Db 2081 CTGCTGATGAGCCACGTCAGCTTGACACAGAAAGCGACGACGTGGTTACAGTGGCT 2140
Qy 122 LeuGluLys 124
Db 2141 CTGGATAAG 2149

RESULT 7
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; APPLICANT: Robinson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; USING IMMUNOLOGICAL AGENTS SPECIFIC FOR P-GLYCOPROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1
Alignment Scores:
Pred. No.: 1,38e-53 Length: 4669
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
Gaps: 4
```

```
US-09-873-409-3 (1-131) x US-09-316-167-1 (1-4669)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTTGGTGTG 1840
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1841 GTGAGTCAGGAACCTGTATTGTTGCCACCGATAGCTGAAACAAATTCGCTATGGCCGT 1900
Qy 42 AspAspValThrAspGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPhe 61
Db 1901 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATCGCTATGACTTT 1960
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1961 ATCATGAAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGAGAGGGCCCAAGTTCAGT 2020
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 2021 GGTGGGCAAGCAGAGGATCGCATTCGACGCGCCCTGTTGCGAACCCCAAGATCCTC 2080
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerIysSerAlaValGlnAlaA 121
Db 2081 CTGCTGATGAGCCACGTCAGCTTGACACCGCTGTTGCGAACCCCAAGATCCTC 2140
Qy 122 LeuGluLys 124
Db 2141 CTGGATAAG 2149

RESULT 8
5206352-3
; Patent No. 5206352
; APPLICANT: Robinson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3
Alignment Scores:
Pred. No.: 1,38e-53 Length: 4669
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
Gaps: 6
```



```
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Alignment Scores:
Pred. No.: 3 7e-53 Length: 9318
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 2 Gaps: 0

US-09-873-409-3 (1-131) x US-08-793-610-6 (1-9318)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 3132 GTTGATGGACAGAGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAATCATTTGGTGTG 3191
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 3192 GTGAGTCAGGAACCTGTATTGTTTGGCCACCAGCATAGCTGAAAACATTCGCTATGGCCGT 3251
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 3252 GAAATGTCAACCATGGATGATGAGTAAGTGTGTAAGAGCTGTCAAGGAAGCAATGCTATGACTTT 3311
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 3312 ATCATGAACTGCTCATATAATTTGACACCTGTGTTGAGAGAGAGGGGCCAGTTGAGT 3371
QY 82 GlyGlnGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 3372 GGTGGGCGAAGCAGAGGATCGCCATTCACGTGCGCCCTGGTTCGCAACCCCAAGATCCTC 3431
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
Db 3432 CTGCTGGATGAGGCCACGTGAGCCTTTGGACACAGAAAGCGAGCGTGTTCAGGTGGCT 3491
QY 122 LeuGluLys 124
Db 3492 CTGGATAAG 3500

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Alignment Scores:
Pred. No.: 3 7e-53 Length: 9318
Score: 461.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 69.95% Indels: 0
DB: 2 Gaps: 0

US-09-873-409-3 (1-131) x US-08-793-610-6 (1-9318)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 3132 GTTGATGGACAGAGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAATCATTTGGTGTG 3191
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 3192 GTGAGTCAGGAACCTGTATTGTTTGGCCACCAGCATAGCTGAAAACATTCGCTATGGCCGT 3251
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 3252 GAAATGTCAACCATGGATGATGAGTAAGTGTGTAAGAGCTGTCAAGGAAGCAATGCTATGACTTT 3311
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 3312 ATCATGAACTGCTCATATAATTTGACACCTGTGTTGAGAGAGAGGGGCCAGTTGAGT 3371
QY 82 GlyGlnGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 3372 GGTGGGCGAAGCAGAGGATCGCCATTCACGTGCGCCCTGGTTCGCAACCCCAAGATCCTC 3431
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121
Db 3432 CTGCTGGATGAGGCCACGTGAGCCTTTGGACACAGAAAGCGAGCGTGTTCAGGTGGCT 3491
QY 122 LeuGluLys 124
Db 3492 CTGGATAAG 3500

RESULT 12
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
```

```
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ PATENT NO. 592898
/ GENERAL INFORMATION:
/ APPLICANT: Skatrud, Paul L.
/ APPLICANT: de Waard, Maarten A.
/ APPLICANT: Peery, Robert B.
/ APPLICANT: Andrade, Alan C.
/ TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
/ TITLE OF INVENTION: Aspergillus nidulans
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eli Lilly and Company
/ STREET: Lilly Corporate Center
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: U.S.
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/461,823
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/168,621
/ FILING DATE: 16-DEC-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/010,433
/ FILING DATE: 27-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Karta, Glenn E.
/ REGISTRATION NUMBER: 30,649
/ REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 301 527-2058
/ TELEFAX: 301 208-6997
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2726 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ US-08-461-823-1

Alignment Scores:
Pred. No.: 1.33e-41 Length: 2726
Score: 372.00 Matches: 75
Percent Similarity: 77.78% Conservative: 23
Best Local Similarity: 59.52% Mismatches: 26
Query Match: 56.45% Indels: 2
DB: 1 Gaps: 1

US-09-873-409-3 (1-131) x US-08-461-823-1 (1-2726)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 1787 CTGCTTGTATGGCAAGAAATAAAGCGACTGATGTTGAGTGGCTCCGAGCACACCTGGGC 1846
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 1847 ATCGTGTCCAGGAGCCCATCTGTTGACTGCAGCATTTGCTGAGAACATTCCTATGGA 1906
QY 41 ArgAspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
Db 1907 GACAAAGCCCGGGTGGTGTACAGAGAGATGCTGAGGCGCAGCAAGAGGAGGCGCAACATA 1966
QY 59 TyrAspPheIleMetGluPhePheAsnThrLeuValGlyGluLysGlyAla 78
Db 1967 CATGCCCTTATCGAGTCACTGCCCTTAATAATATAGCACTAAGTAGAGACAAAGGAACT 2026
QY 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
Db 2027 CAGCTCTCTGTGGCCAGAAACACGACATGTCATGCTGTCGCTTGTGTAGACAGCCT 2086
QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
Db 2087 CATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTC 2146
QY 119 GlnAlaAlaLeuGluLys 124
Db 2147 CAGAAGCCCTTGGACAAA 2164
RESULT 13
US-08-996-545-1

/ Sequence 1, Application US/08996545
/ Patent No. 592898
/ GENERAL INFORMATION:
/ APPLICANT: Skatrud, Paul L.
/ APPLICANT: de Waard, Maarten A.
/ APPLICANT: Peery, Robert B.
/ APPLICANT: Andrade, Alan C.
/ TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
/ TITLE OF INVENTION: Aspergillus nidulans
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eli Lilly and Company
/ STREET: Lilly Corporate Center
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: U.S.
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/996,545
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Webster, Thomas D.
/ REGISTRATION NUMBER: 39,872
/ REFERENCE/DOCKET NUMBER: X-11766
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-276-3334
/ TELEFAX: 317-276-2763
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4002 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..4002
/ US-08-996-545-1

Alignment Scores:
Pred. No.: 1.56e-40 Length: 4002
Score: 366.00 Matches: 74
Percent Similarity: 73.60% Conservative: 18
Best Local Similarity: 59.20% Mismatches: 31
Query Match: 55.54% Indels: 2
DB: 2 Gaps: 1

US-09-873-409-3 (1-131) x US-08-996-545-1 (1-4002)
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 3445 CTTGTTGATGGGAGGACATAGTAACATAATATCACTTACCAGCAGCTTCTGTGCA 3504
QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db 3505 CTGCTCAGCAGAGCGGACACTGTACCGGCGCCATCAAGGAAACATCTTACTTGGT 3564
QY 41 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
Db 3565 ATTGTCGAGATGACCTACCGGAGAAATCTTGTATTAGGCTTGCAAGGACGCTAATATC 3624
QY 59 TyrAspPheIleMetGluPhePheAsnThrLeuValGlyGluLysGlyAla 78
Db 3625 TAGGACTTCATCATGCTCGCTCCCGAGGGGCTTTAATACAGTTGTTGGCAAGGAGGC 3684
```

QY 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsnPro 98  
Db 3685 ATGTTCTTGGGGCCCAAGCAACGTGTGGCCATTGCGGAGCCCTTCTTGGGGATCCC 3744  
QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118  
Db 3745 AAAATCTTCTTCTCGATGAGGCGACGTAGCCCTCGACTCCGAGTCAGAAAAGGTGCTC 3804  
QY 119 GlnAlaAlaLeuGlu 123  
Db 3805 CAGCGGCTTTGGAT 3819  
RESULT 14  
US-08-996-545-3  
; Sequence 3, Application US/08996545  
; Patent No. 5928898  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene at:rD of  
; TITLE OF INVENTION: Aspergillus Nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/996,545  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4002 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-996-545-3  
Alignment Scores:  
Pred. No.: 1.56e-40 Length: 4002  
Score: 366.00 Matches: 74  
Percent Similarity: 73.60% Conservative: 18  
Best Local Similarity: 59.20% Mismatches: 31  
Query Match: 55.54% Indels: 2  
Db: 2 Gaps: 1  
US-09-873-409-3 (1-131) x US-08-996-545-3 (1-4002)  
QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
Db 3445 CUUGUUGAGGGAGGACAAAGCAACUAAUUAUACUCCUCCGAGCAGCUUUGUCA 3504  
QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleTyrGly 40

Db 3505 CUGGUCAGCCAGGAGCGGACACUGUACCGAGGGCACCACUAAAGAAACAUCUUAUUGGU 3564  
QY 41 -----ArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58  
Db 3565 AUUGUGAAGAGUAGCUGACCGGAGAAUUCUUAUAGGCUUGCAAGGACGACGCUAAUUC 3624  
QY 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyAla 78  
Db 3625 UACGACUUAUAGUUGCGUCCCGGAGGCUUUAUACAGUUGUUGGACGACGAGGAGGC 3684  
QY 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsnPro 98  
Db 3685 AUGUUGUCUGGCGGCAAGCAACGUGUGGCAUUGCCGAGCCUUCUUGGGAUCCC 3744  
QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118  
Db 3745 AAAAUUU 3804  
QY 119 GlnAlaAlaLeuGlu 123  
Db 3805 CAGCGGCUUUGGAU 3819  
RESULT 15  
US-09-328-320-1  
; Sequence 1, Application US/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene at:rD of  
; TITLE OF INVENTION: Aspergillus Nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/328,320  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: 08/996,545  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4002 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..4002

US-09-328-320-1

Alignment Scores:

Pred. No.:	1.56e-40	Length:	4002
Score:	366.00	Matches:	74
Percent Similarity:	73.60%	Conservative:	18
Best Local Similarity:	59.20%	Mismatches:	31
Query Match:	55.54%	Indels:	2
DB:	4	Gaps:	1

US-09-873-409-3 (1-131) x US-09-328-320-1 (1-4002)

Qy	1	MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly	20
Db	3445	CTTGTTCATGGGAAGGACATAAGTAAATATCACTCACTACCGCAGCTTCTGTCA	3504
Qy	21	ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly	40
Db	3505	CTGGTCAGCCAGGAGCGCACACTGTACCGGGCACCATCAAGGAAACATCTTACTTGGT	3564
Qy	41	-----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla	58
Db	3565	ATTGTCGAAGATGACGTACCGGAGAGATTCTTGATTAAGCTTGCAGGACGCTAATATC	3624
Qy	59	TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla	78
Db	3625	TACGACTTCATCATGTCGCTCCCGAGGGGCTTTAATACAGTTGTTGGCAGCAAGGAGGC	3684
Qy	79	GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro	98
Db	3685	ATGTTGTCTGGCGGCCAAAGCAACGTCGTGCCCATTTGCCGAGCCCTTCTTCGGGATCC	3744
Qy	99	LysIleIleuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal	118
Db	3745	AAATCCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAGGTCGTC	3804
Qy	119	GlnAlaAlaLeuGlu	123
Db	3805	CAGGCGCTTTGGAT	3819

Search completed: March 31, 2003, 15:23:26  
Job time : 39.1843 secs

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 901.042 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-3  
Perfect score: 659  
Sequence: 1 MVDENDIRALNVRHYRDHIG.....SESKSAVQALEKDTFRYSF 131

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US9873409/runat\_27032003\_115420\_19240/app\_query.fasta\_1.7544  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US9873409 @CGN 1 1 30544 @runat\_27032003\_115420\_19240 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pin.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	618	93.8	760	13	BM016204	BM016204 603642659
2	578	87.7	559	12	BF692596	BF692596 602248949
3	524	79.5	545	9	AL040762	AL040762 DXFZ9434C
4	490	74.4	669	13	BM013981	BM013981 603639619
5	464.5	70.5	1033	14	BM926413	BM926413 AGENCOURT
6	445	67.5	550	13	BJ075336	BJ075336 BJ075336
7	428	64.9	1098	17	CNS051EN	AL316760 Tetraodon
8	393	59.6	550	13	BJ388678	BJ388678 BJ388678
9	379	57.5	520	13	BJ336809	BJ336809 BJ336809
10	378.5	57.4	589	9	AI666987	AI666987 fc24c04.y
11	378.5	57.4	589	9	AI666988	AI666988 fc24c04.y
12	377	57.2	542	13	BJ339421	BJ339421 BJ339421
13	377	57.2	547	13	BJ333166	BJ333166 BJ333166
14	374	56.8	589	13	BJ356968	BJ356968 BJ356968
15	374	56.8	608	13	BJ356648	BJ356648 BJ356648
16	374	56.8	656	13	BJ401771	BJ401771 BJ401771
17	374	56.8	687	13	BJ350425	BJ350425 BJ350425
18	374	56.8	687	13	BJ375176	BJ375176 BJ375176
19	374	56.8	688	13	BJ405288	BJ405288 BJ405288
20	374	56.8	690	13	BJ408755	BJ408755 BJ408755
21	374	56.8	692	13	BJ407928	BJ407928 BJ407928
22	374	56.8	696	13	BJ348083	BJ348083 BJ348083
23	374	56.8	722	13	BJ406558	BJ406558 BJ406558
24	374	56.8	731	13	BJ354552	BJ354552 BJ354552
25	374	56.8	733	13	BJ376217	BJ376217 BJ376217
26	374	56.8	739	13	BJ404475	BJ404475 BJ404475
27	374	56.8	741	13	BJ347139	BJ347139 BJ347139
28	374	56.8	742	13	BJ372910	BJ372910 BJ372910
29	374	56.8	743	13	BJ408863	BJ408863 BJ408863
30	374	56.8	743	13	BJ405178	BJ405178 BJ405178
31	374	56.8	758	13	BJ357956	BJ357956 BJ357956
32	374	56.8	783	13	BJ352760	BJ352760 BJ352760
33	374	56.8	784	13	BJ405287	BJ405287 BJ405287
34	374	56.8	792	13	BJ352404	BJ352404 BJ352404
35	374	56.8	814	13	BJ352304	BJ352304 BJ352304
36	374	56.8	820	13	BJ356540	BJ356540 BJ356540
37	374	56.8	835	13	BJ349604	BJ349604 BJ349604
38	374	56.8	838	13	BJ355307	BJ355307 BJ355307
39	372	56.4	943	9	AL520322	AL520322 AL520322
40	370	56.1	681	14	BU006064	BU006064 QGG9N16.y
41	370	56.1	715	14	BQ869512	BQ869512 QGD6G09.y
42	370	56.1	894	12	BF584668	BF584668 602098406
43	369	56.0	508	14	BQ818890	BQ818890 1030074A0
44	369	56.0	635	13	BI716519	BI716519 1031010G1
45	368.5	55.9	871	9	AL582625	AL582625 AL582625

ALIGNMENTS

RESULT 1  
BM016204  
LOCUS 603642659F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5418615 5', mRNA linear EST 30-OCT-2001  
DEFINITION mRNA sequence.  
ACCESSION BM016204.1 GI:16530558  
VERSION BM016204.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 760)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)





```

Db 326 GATTTTAGATGAGGCTACCTCTGCGCCCTGGATTACAGAAAGGAGTCAGCTGTTCAAGCTGC 385
|||||
Qy 121 aleuGluLysAspThrProArgTyrSerPhe 131
|||||
Db 386 ACTGGAGAGGATACCCCGAGGTATTCATT 416
|||||
RESULT 3
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKFZp434C1815 r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKFZp434C1815 5', mRNA sequence.
ACCESSION AL040762
VERSION AL040762.1 GI:5409708
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferseitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKFZp434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN
Alignment Scores:
Pred. No.: 1,76e-60 Length: 545
Score: 524.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.51% Indels: 0
DB: 9 Gaps: 0
US-09-873-409-3 (1-131) x AL040762 (1-545)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
|||||
Db 237 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGGCGCATTCAGACCATATATGGA 296
|||||
Qy 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
|||||
Db 297 GTGGTTAGTCAAGAGCGCTGTTTGTTCGGACCAACCATCAGTAACCAATATCAAGTATGGA 356
|||||
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaLeuValArgGluAlaAsnAlaTyrAsp 60
|||||
Db 357 CGAGATGATGTGCTGATGAAGAGATGGAGAGCAGCAGGAGGAGCAATATGCTGATGAT 416
|||||
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
|||||

```

```

Db 417 TTTATCATGGAGTTTCCCTAAATAATTTTAAATCATTTGGTAGGGAAAAGAGACTCAATG 476
|||||
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
|||||
Db 477 ACTGGAGGGCAGAAACAGAGGATCGCAATTGCTGTCGCTTAGTTCGAAACCCCAAGATT 536
|||||
Qy 101 LeuIleLeu 103
|||||
Db 537 CTGATTTTA 545
|||||
RESULT 4
EM013981
LOCUS 60363961F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
DEFINITION mRNA sequence.
ACCESSION BM013981
VERSION BM013981.1 GI:16528335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2060 row: f column: 05
High quality sequence stop: 664.
Location/Qualifiers
source
1..669
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5415556"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site_2: SalI; Cloned unidirectionally, oligo-dT primed.
Average insert size 1.383 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 203 a 122 c 162 g 182 t
ORIGIN
Alignment Scores:
Pred. No.: 9,06e-56 Length: 669
Score: 490.00 Matches: 98
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 0
Query Match: 74.36% Indels: 1
DB: 13 Gaps: 0
US-09-873-409-3 (1-131) x BM013981 (1-669)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
|||||
Db 374 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGGCGCATTCAGACCATATATGGA 433
|||||
Qy 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
|||||
Db 434 GTGGTTAGTCAAGAGCGCTGTTTGTTCGGACCAACCATCAGTAACCAATATCAAGTATGGA 493
|||||
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaLeuValArgGluAlaAsnAlaTyrAsp 60
|||||

```

```

Db 494 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 553
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 554 TTTTATCATGAGATTTCTTAATAAATTAATACATTTGGTAGGGGAAAAGAGAGCTCAAAATG 613
Qy 81 SerGlyGlyGlnLysGlnArg-1leAlaIleAlaArgAlaLeuValArgAsnPro 98
Db 614 AGTGGAGGCGAGAAACAGAGGCATCGCAATTTGCTGTGCTTGAATTTGCAAAACCC 668

RESULT 5
BM926413
LOCUS BM926413
DEFINITION AGENCOURT 6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
5', mRNA sequence.
ACCESSION BM926413
VERSION BM926413.1 GI:19376792
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1033)
AUTHORS NIH-MGC http://img.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapps@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12818 row: o column: 22
High quality sequence start: 3
High quality sequence stop: 686.
Location/Qualifiers
1. .1033
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5764845"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 289 a 212 c 218 g 312 t
ORIGIN
source

Alignment Scores:
Pred. No.: 4,54e-52 Length: 1033
Score: 464.50 Matches: 111
Percent Similarity: 87.02% Conservative: 3
Best Local Similarity: 84.73% Mismatches: 17
Query Match: 70.49% Indels: 4
DB: 14 Gaps: 0

US-09-873-409-3 (1-131) x BM926413 (1-1033)
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db 534 ATGGTGGATGAGATGATCATCAGACGCTTTAAATGTGGCGCATATATGAGACCAATATGGA 593
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40

```

```

Db 594 GTGGTTAGTCAAGAGCCTGTTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 653
Qy 41 ArgAspAspValThrAspGluMetGluArgAlaAlaAargGluAlaAsnAlaTyrAsp 60
Db 654 CGAGATGATGTGCTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAATGCGTATGAT 713
Qy 61 PheileMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
Db 714 TTTATCATGAGATTTCTTAATAA-TTATATACATTTGGTAGGGGAAAAGAGAGCTCCAATG 772
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaAargAlaLeuValArgAsnProLysIle 100
Db 773 AGTGGAGGCGAGAAACCG-AGGATCCCAATTGCTCTGCTTGAATTCGAAACCCAGGAT- 830
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
Db 831 CTGAATTTAGATGAAGCTACGCTCTGCCCTGGATTTCAGAAACCAAGTCAACTGGTCA--GCT 888
Qy 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
Db 889 GGCTGGAGAGGAAACCCCGGNAATTCCTTT 921

RESULT 6
BM925336
LOCUS BM925336
DEFINITION laevis cDNA clone XL057d10 5', mRNA sequence.
ACCESSION BM925336
VERSION BM925336.1 GI:17520252
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 550)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .550
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL057d10"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue type="whole embryo"
/dev stage="stage 25"
/notes="Vector: PBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 162 a 109 c 144 g 135 t
ORIGIN
source

Alignment Scores:
Pred. No.: 8,26e-50 Length: 550
Score: 445.00 Matches: 86
Percent Similarity: 90.00% Conservative: 22
Best Local Similarity: 71.67% Mismatches: 12
Query Match: 67.53% Indels: 0
DB: 13 Gaps: 0

```

US-09-873-409-3 (1-131) x BJ075336 (1-550)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 191 TTGGATGGACACCATATCCGTTCACTAAATGTAATAATGGTTGAGGGAAAATATTCGCGTA 250  
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 251 GTGAGTCAGGAGCAGTCGCTTTTGGCACAAACATTCGAGAAATATCCGCTACGCGAGG 310  
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 Db 311 GAATCAGTAACAGACACGAAATGTAACAGCGGCAAGAGCAATGCTTTGATTTT 370  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 371 ATATCCAGGCTTCCTGATAAATTTAAGACGATGGTGGGAGAAAGAGTGGCGCAGTTGAGT 430  
 QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 431 GGGGGTCAGAAGCAGCGGATAGCCATTCGACGCGCCCTGGTACGGAACCCCTAAGATTCTA 490  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 491 CTCTTGATGAACCCACATCAGCTCTGGATACACAAAGTGAAGCTATTGTTACAAAGCTGCC 550

# RESULT 7 CNS051EN/c

LOCUS 1098 bp DNA linear GSS 26-JUL-2000  
 DEFINITION Tetraodon nigroviridis genome survey sequence Y7 end of clone  
 052D12 of library A from Tetraodon nigroviridis, genomic survey  
 sequence.

ACCESSION AL316760  
 VERSION AL316760.1 GI:9549644

KEYWORDS GSS; genome survey sequence.  
 SOURCE Tetraodon nigroviridis.

ORGANISM Tetraodon nigroviridis.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 1098)  
 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,  
 Saurin,W. and Weissenbach,J.  
 Estimate of human gene number provided by genome-wide analysis  
 using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE  
 PUBMED 20296633

REFERENCE 2 (bases 1 to 1098)  
 Crolius,H.R., Jaillon,O., Dasilva,C., Ozouf-Coataz,C., Fizames,C.,  
 Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W.,  
 Bernot,A. and Weissenbach,J.

TITLE Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE  
 PUBMED 10899143

REFERENCE 3 (bases 1 to 1098)  
 Genoscope.  
 Direct Submission

TITLE Submitted (12-APR-2000)

COMMENT This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the tetraodon nigroviridis  
 genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/Tetraodon.

## FEATURES source

1..1098  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="052D12"  
 /clone\_lib="A"

BASE COUNT 232 a 281 c 312 g 260 t 13 others  
 ORIGIN /note="Genoscope sequence ID : C0AA052DB06C1-end : T7"

Alignment Scores:  
 Pred. No.: 4,19e-47 Length: 1098  
 Score: 428.00 Matches: 89  
 Percent Similarity: 75.34% Conservative: 21  
 Best Local Similarity: 60.96% Mismatches: 13  
 Query Match: 64.95% Indels: 23  
 DB: 17 Gaps: 1

US-09-873-409-3 (1-131) x CNS051EN (1-1098)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1005 ATCGACGGTCACGACATCCGCTCTCTCAACATCCGCTACCTGAGAGAGATGATCGGGTG 946

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 945 GTGAGTCAGAGCCCATCCTCTTCGCCACCCATCGCCGAGAACATCCGATACGCCGA 886

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 Db 885 CTGGAGCTGACCGAGGAGATCGAACGAGCCACCAAGAGTCCACGCTTACGACTTC 826

QY 62 IleMetGluPheProAsn----- 67  
 Db 825 ATCATGAGCCTTCCAGATGTGAGAGCCGGTTGGTCCCCCGCTCGGTCCACCCCTGGCTGA 766

QY 68 -----LysPheAsnThrLeuValGlyGluLysGlyAlaG 79  
 Db 765 TCTCTGCTCCGTCGTCCTCTGAGCAGAGAGTTTGAGACGCTGGTGGGAGACCGCGGACTCA 706

QY 79 nMetSerGlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLys 99  
 Db 705 GTTGACGGGAGGACAGAACGAGGATCCGATCCGATCCGATCCGATCCGATCCGATCCG 646

QY 99 sIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 119  
 Db 645 ATCTCTGCTGCTGGATGAAGCCACGTCGCTCGCTGGATGCTGAGAGCGAGACCATCTACA 586

QY 119 nAlaAlaLeuGluLys 124  
 Db 585 GCGTCTCTGGCAAG 570

RESULT 8  
 BJ388678

LOCUS 550 bp mRNA linear EST 08-MAR-2002  
 DEFINITION Dictyostelium discoideum cDNA library, SF Dictyostelium  
 discoideum cDNA clone dds6p15 5', mRNA sequence.

ACCESSION BJ388678  
 VERSION BJ388678.1 GI:19299764

KEYWORDS EST.  
 SOURCE Dictyostelium discoideum.

ORGANISM Dictyostelium discoideum  
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE 1 (bases 1 to 550)  
 Urushinara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

AUTHORS Full length cDNA of Dictyostelium discoideum at the slug stage  
 TITLE Unpublished (2002)  
 JOURNAL Contact: Tadao Shin-i

COMMENT Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshini@genes.nig.ac.jp.

FEATURES  
 source 1..550  
 /organism="Dictyostelium discoideum"  
 /strain="AX4"  
 /db\_xref="taxon:44689"

```

/clone="ddse6p15"
/clone_lib="Dictyostelium discoideum cDNA library, SP"
/sex="mat A"
/dev_stage="Slug stage"
BASE COUNT 190 a -77 c 113 g 170 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 8.66e-43 Length: 550
Score: 393.00 Matches: 76
Percent Similarity: 78.86% Conservative: 21
Best Local Similarity: 61.79% Mismatches: 26
Query Match: 59.64% Indels: 0
DB: 13 Gaps: 0

```

US-09-873-409-3 (1-131) x BJ388678 (1-550)

```

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 96 TTATTGGATGGAGAGATATTAGAAAATTCATATGTCAGAGGGTTACGTCAAAATAATGGT 155

```

```

QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
DB 156 TTAGTCAATCAAGAGCGCTGTATTATTCGCCCAACCATTAGTGAATAATATTCGTTATGGT 215

```

```

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 216 AAGAAGGTCCTCAAGATGAATTAAGAGAGCGAGCTAAATAGCAATGCCATAGT 275

```

```

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 276 TTTATATCACAATTACCAACAGCGGTATATACTTTGTTAGTGCAGAAAGGTGCCAAATG 335

```

```

QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 336 TCCTGGTGGTCAAGACCAACGTATAGCAATTCGTCGCGCTTTATTAATAATCCAAATATA 395

```

```

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 396 TTACTCTTGATGAATCAACATCTGCGATTCGATGCAGAGTCTACAAATAGTACAAAGAG 455

```

```

QY 121 AlaLeuGlu 123
DB 456 GCATTAGAT 464

```

```

RESULT 9
BJ336809 BJ336809 520 bp mRNA linear EST 05-MAR-2002
LOCUS BJ336809 Dictyostelium discoideum cDNA library, AF Dictyostelium
DEFINITION discoideum cDNA clone dda55p03 5', mRNA sequence.
ACCESSION BJ336809
VERSION BJ336809.1 GI:19166939
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum

```

```

REFERENCE 1 (bases 1 to 520)
AUTHORS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
TITLES Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
Full length cDNA of Dictyostelium discoideum at the aggregation
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. 520
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"

```

```

FEATURES
source
1. 520
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"

```

```

/clone="dda55p03"
/clone_lib="Dictyostelium discoideum cDNA library, AF"
/sex="mat A"
/dev_stage="Aggregation stage"
BASE COUNT 178 a -78 c 106 g 158 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 6.23e-41 Length: 520
Score: 379.00 Matches: 73
Percent Similarity: 78.81% Conservative: 20
Best Local Similarity: 61.86% Mismatches: 25
Query Match: 57.51% Indels: 0
DB: 13 Gaps: 0

```

US-09-873-409-3 (1-131) x BJ336809 (1-520)

```

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 165 TTATTGGATGGAGAGATATTAGAAAATTCATATGTCAGAGGGTTACGTCAAAATAATGGT 224

```

```

QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGly 40
DB 225 TTAGTCAATCAAGAGCGCTGTATTATTCGCCCAACCATTAGTGAATAATATTCGTTATGGT 284

```

```

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 285 AAGAAGGTCCTCAAGATGAATTAAGAGAGCGAGCTAAATAGCAATGCCATAGT 344

```

```

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 345 TTTATATCACAATTACCAACAGCGGTATATACTTTGTTAGTGCAGAAAGGTGCCAAATG 404

```

```

QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
DB 405 TCCTGGTGGTCAAGACCAACGTATAGCAATTCGTCGCGCTTTATTAATAATCCAAATATA 464

```

```

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
DB 465 TTACTCTTGATGAATCAACATCTGCGATTCGATGCAGAGTCTACAAATAATAGTA 518

```

```

RESULT 10
A1666987 588 bp mRNA linear EST 07-JUN-2001
LOCUS fc24c03.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
DEFINITION IMAGE:3722308 5', similar to TR:Q13040 Q13040 ATP-BINDING CASSETTE
PROTEIN ;, mRNA sequence.
ACCESSION A1666987
VERSION A1666987.1 GI:4805343
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 588)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,L., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rietter,E.,
Kohn,S., Shin.T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@wustl.edu
CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,

```

Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)

Seq primer: T3 ET from Amersham  
High quality sequence stop: 481.  
Location/Qualifiers

#### FEATURES

source

1..588  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="IMAGE:3722308"  
/clone\_lib="Zebrafish Washu MPIMG EST"  
/sex="mixed"  
/tissue type="26 somite embryos, adult livers, shield stage embryos"  
/lab host="XLI-blue MRF"  
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer 5'-pGACTAGTCTAGATCGGAGCGGCCCTTTTITTTT3'; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 149 a 138 c 167 g 132 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 8.61e-41 Length: 588  
Score: 378.50 Matches: 79  
Percent Similarity: 77.78% Conservative: 19  
Best Local Similarity: 62.70% Mismatches: 25  
Query Match: 57.44% Indels: 3  
DB: 9 Gaps: 1

US-09-873-409-3 (1-131) x AI666987 (1-588)

Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
Db 1 GTTGATGGTCATGATATTCGAGACTTGAATCCCTTACTGGTTAGGAGTCACATGGTACT 60  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleYsTyrGlyArg 41  
Db 61 GTAAGCCAGGAACCGGTCTCTTCTCTGTTCAATTGCTGAGAACATTCGTTATGAGCG 120  
Qy 42 AspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58  
Db 121 CCGATCCGAGTCAGTGACAGTGGAGGACATTCAGGCGGCACAGATCGCAATGCT 180  
Qy 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyAla 78  
Db 181 CATGACTTCATCCAGGAGTTTCCCAAGGGGTACGACTACTGTGTGGGAGAGAGGGGTT 240  
Qy 79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98  
Db 241 CTGCTGTGAGTGGCAAAAGCAGAGAGTGGCCATGACGAGCCCTGCTGGAAGATCCC 300  
Qy 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118  
Db 301 AAATACTTTTCTGATGATGAGCCAGGAGTGTCTGTGATGACAGATGATGATGTTTGGTC 360  
Qy 119 GlnAlaLeuGlyLys 124

Db 361 CAGGAGCGCTTAGAAAGG 378  
RESULT 11  
AI666988

LOCUS  
DEFINITION

AI666988 589 bp mRNA linear EST 07-JUN-2001  
fc24c04.y1 Zebrafish Washu MPIMG EST Danio rerio cDNA clone  
IMAGE:3722310 5' similar to TR:Q13040 Q13040 ATP-BINDING CASSETTE  
PROTEIN ; mRNA sequence.

ACCESSION AI666988  
VERSION AI666988.1 GI:4805344  
KEYWORDS EST.  
SOURCE zebrafish.  
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.

REFERENCE  
AUTHORS

1 (bases 1 to 589)  
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy  
S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood  
K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,  
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,  
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.  
and Wilson,R.  
Washu Zebrafish EST Project 1998

TITLE  
JOURNAL  
COMMENT

Unpublished (1998)  
Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrafish@watson.wustl.edu

cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:  
Matthew Clark. DNA Sequencing by: Washington University Genome  
Sequencing Center Clone distribution: Genome Systems, St. Louis,  
Missouri (web address: www.genomesystems.com) (email contact:  
info@genomesystems.com) and Research Genetics, Huntsville, Alabama  
(web address: www.resgen.com) (email contact: info@resgen.com) and  
RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address:  
www.rzpd.de)

Seq primer: T3 ET from Amersham  
Seq quality sequence stop: 481.  
Location/Qualifiers

1..589  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone="IMAGE:3722310"  
/clone\_lib="Zebrafish Washu MPIMG EST"  
/sex="mixed"  
/tissue type="26 somite embryos, adult livers, shield stage embryos"  
/lab host="XLI-blue MRF"  
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer 5'-pGACTAGTCTAGATCGGAGCGGCCCTTTTITTTT3'; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 151 a 137 c 167 g 133 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 8.63e-41 Length: 589  
 Score: 378.50 Matches: 79  
 Percent Similarity: 77.78% Conservative: 19  
 Best Local Similarity: 62.70% Mismatches: 25  
 Query Match: 57.44% Indels: 3  
 DB: 9 Gaps: 1

US-09-873-409-3 (1-131) x AI666988 (1-589)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 DB 1 GTTGATGGTCAATGATATTCAGAGCTTGAATCCTATTGGTTAAGGAGTCACATGGTACT 60  
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 DB 61 GTAAGCCAGGACCGGTCCTTTCTCTGTTCAATTCCTGAGAACATTCGTTATGAGCG 120  
 QY 42 AspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58  
 DB 121 CCGATCCGAGTCAGGTGACAGTGGAGGACATTTACAGGGCGGCACAGATCGCCAAATGCT 180  
 QY 59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyValAla 78  
 DB 181 CATGACTTCATCCAGGAGTTCCTCAAGGGCTACGATACCTGTGTGGAGAGAGGGGGT 240  
 QY 79 GlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98  
 DB 241 CTGCTGTGAGTGAGCAAAAGCAGAGAGTGGCCATAGCAGAGCCCTGCTGAGAGATCCC 300  
 QY 99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118  
 DB 301 AAAATACTTTTACTGGATGAAGCCAGGTCCTCTGGATGACAGAGATGAGTTTGGTC 360  
 QY 119 GlnAlaLeuGluLys 124  
 DB 361 CAGGAGCGGTTAGAAAGG 378

## RESULT 12

BJ339421

LOCUS

DEFINITION

BJ339421 Dictyostelium discoideum cDNA library, AF Dictyostelium

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

UNPUBLISHED (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp

Location/Qualifiers

FEATURES

source

1. .542

/organism="Dictyostelium discoideum"

/strain="AX4"

/db\_xref="taxon:44689"

/clone="dda65d11"

/sex="mat A"

/dev\_stage="Aggregation stage"

BASE COUNT 186 a 77 c 113 g 166 t

ORIGIN

## Alignment Scores:

Pred. No.: 1.23e-40 Length: 542  
 Score: 377.00 Matches: 72  
 Percent Similarity: 80.70% Conservative: 20  
 Best Local Similarity: 63.16% Mismatches: 22  
 Query Match: 57.21% Indels: 0  
 DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x BJ339421 (1-542)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
 DB 192 TTATTCGATGGAGAGAGATATTAGAAAATTCATGTGAGAGGGTTAGCTCAAAAATTTGGT 251  
 QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
 DB 252 TTAGTCGAATCAAGAGCTGTATTATTCGCCACACACCATTTAGTCAAAATATTCGTTATGGT 311  
 QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
 DB 312 AAAGAAGGTGCCACTCAGATGAAATTTGAAGAAGCAGCTAAATTTAGCAATGCCATAGT 371  
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMet 80  
 DB 372 TTATATCAATCAATTCACAGGGTTATAATCTTCGTAGGTGAGAAAGGTGTCCAAATG 431  
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
 DB 432 TCTGGTGGTCAAGACACAGTATAGCAATTCGTGCGCTTATTAAAAATCCAAATATA 491  
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 114  
 DB 492 TTACTCTTGATGAATCAACATCTGCAATTCGATTCGATGACAGTCT 533

## RESULT 13

BJ333166

LOCUS

DEFINITION

BJ333166 Dictyostelium discoideum cDNA library, AF Dictyostelium

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

UNPUBLISHED (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp

Location/Qualifiers

FEATURES

source

1. .547

/organism="Dictyostelium discoideum"

/strain="AX4"

/db\_xref="taxon:44689"

/clone="dda41m18"

/sex="mat A"

/dev\_stage="Aggregation stage"

BASE COUNT 186 a 78 c 114 g 169 t

ORIGIN

## Alignment Scores:

Pred. No.: 1.24e-40 Length: 547  
 Score: 377.00 Matches: 72  
 Percent Similarity: 80.70% Conservative: 20



Best Local Similarity: 63.16% Mismatches: 22  
 Query Match: 57.21% Indels: 0  
 DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x BJ333166 (1-547)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisThrArgAspHisIleGly 20  
 DB 195 TTTATGGATGGGAAGATATTAGAAAATTCAATGTGCAGAGGGTTAGCTCAAAAATTTGGT 254  
 QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysIleGly 40  
 DB 255 TTATGCAATCAAGAGCTGTATTATTTGCCACACCAATTTAGTCAAAATATTGTTATGGT 314  
 QY 41 ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaIleThrAsp 60  
 DB 315 AAAGAAGGTGCCACTCAAGATGAATTAAGAGAGCAGCTAAATTTAGCAAAATGCCCATAGT 374  
 QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
 DB 375 TTTATATCACAATTACACAGGGTTATAATCTTTGTTAGTGAGAAAGGTGTCCAAATG 434  
 QY 81 SerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
 DB 435 TCTGGTGGTCAAGACAAAGTATAGCAATTTGCTGCGCCGCTATTAAATAATCCAAATATA 494  
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 114  
 DB 495 TTACTCTTGATGAATCAACATCTGCATTGGATGCAGAGTCT 536

## RESULT 14

BJ356968/c  
 LOCUS BJ356968 589 bp mRNA linear EST 07-MAR-2002  
 DEFINITION BJ356968 Dictyostelium discoideum cDNA library, AF Dictyostelium  
 discoideum cDNA clone dda62e02 3', mRNA sequence.

ACCESSION BJ356968

VERSION BJ356968.1 GI:19256563

KEYWORDS EST.

SOURCE Dictyostelium discoideum.

ORGANISM Dictyostelium discoideum.

REFERENCE 1 (bases 1 to 589) Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE Full length cDNA of Dictyostelium discoideum at the aggregation stage

## JOURNAL

COMMENT Unpublished (2002)

Contact: Tadaasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tschini@genes.nig.ac.jp.

## FEATURES

source

1..589  
 /organism="Dictyostelium discoideum"  
 /strain="AX4"  
 /db\_xref="taxon:44689"  
 /clone="dda62e02"  
 /clone\_lib="Dictyostelium discoideum cDNA library, AF"  
 /sex="mat A"  
 /dev\_stage="Aggregation stage"

BASE COUNT 182 a 103 c 99 g 205 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 3.5e-40 Length: 589  
 Score: 374.00 Matches: 71  
 Percent Similarity: 79.51% Conservative: 26  
 Best Local Similarity: 58.20% Mismatches: 25  
 Query Match: 56.75% Indels: 0  
 DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x BJ356968 (1-589)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisThrArgAspHisIleGlyVal 21  
 DB 541 ATCGATGGTCATAATATAAAGAAATCTAAATGTTGGTCATCTTAGACATCTCTTTGGTTTA 482

QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysIleGlyArg 41  
 DB 481 GTTGGTCAAGAACCAACTCTATTTCAGTGGTACAAATTTGCCGACAATATTCTGTTATGTAAG 422

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaIleThrAspPhe 61  
 DB 421 CATGATGCCAATCAAGAGAGGATTTGAAGAGGCTTCAAAATTTATCAATTTCTCATTTTC 362

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 DB 361 ATCAATTGATTACCAATATGTTTAAACACAGAGTTGGGTGAGAAATACACTCAATTATCA 302

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 DB 301 GGTGGTCAAAAAACAAGTATAGCAATTTGCTCGTGAATCATTCGTAATCCAAAGATTTTA 242

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaIle 121  
 DB 241 TTATTGGATGAGTCAACCTCTGCATTGGATGCTGATTCAACAAATTTAGTTCAAGAAGCA 182

QY 122 LeuGlu 123

DB 181 CTAGAA 176

## RESULT 15

BJ356648/c

LOCUS BJ356648 608 bp mRNA linear EST 07-MAR-2002

DEFINITION BJ356648 Dictyostelium discoideum cDNA library, AF Dictyostelium

discoideum cDNA clone dda61b02 3', mRNA sequence.

ACCESSION BJ356648

VERSION BJ356648.1 GI:19256243

KEYWORDS EST.

SOURCE Dictyostelium discoideum.

ORGANISM Dictyostelium discoideum.

REFERENCE 1 (bases 1 to 608) Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.

TITLE Full length cDNA of Dictyostelium discoideum at the aggregation stage

JOURNAL Unpublished (2002)

Contact: Tadaasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tschini@genes.nig.ac.jp.

Location/Qualifiers

1..608

/organism="Dictyostelium discoideum"

/strain="AX4"

/db\_xref="taxon:44689"

/clone="dda61b02"

/clone\_lib="Dictyostelium discoideum cDNA library, AF"

/sex="mat A"

/dev\_stage="Aggregation stage"

BASE COUNT 187 a 110 c 101 g 210 t

ORIGIN

## Alignment Scores:

Pred. No.: 3.65e-40 Length: 608  
 Score: 374.00 Matches: 71  
 Percent Similarity: 79.51% Conservative: 26  
 Best Local Similarity: 58.20% Mismatches: 25  
 Query Match: 56.75% Indels: 0  
 DB: 13 Gaps: 0

US-09-873-409-3 (1-131) x BJ356648 (1-608)

```

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 539 ATCGATGGTCATAATAAAGAATCTAAATGTTGTCATCTTAGACATCTCTTGGTTA 480
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 479 GTTGGTCAAGAACCAACTCTATTTCAGTGGTACAAATGCCGACAATATTGTTATGTAAG 420
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 419 CATGATGCACACTCAAGAGGAGATTGAAGAGGCTTCAAAATTATCAAAATTCATTCATTC 360
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 359 ATCAATTGATTTCCAAAATGGTTATAACACAGATTGGTGAGAAATACACTCAATTATCA 300
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 299 GGTGGTCNAAAACACACGTATAGCAATTGCTCGTCAATCATTGTAATCCAAAGATTTA 240
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 239 TTATTGGATGAGTCAACCTCTGCATTGGATGCTGATTCAACAAAATTAGTTCAGAAGCA 180
QY 122 LeuGlu 123
Db 179 CTAGAA 174

```

Search completed: March 31, 2003, 13:50:03  
Job time : 907.167 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 128.031 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-3

Perfect score: 659

Sequence: 1 MVDENDIRALNVRHYRDHIG.....SESKSAVQAALKEKTPRYSF 131

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xmlh

-MODEL=frame+ p2n.model -DEV-xmlh  
-Q/cn2\_1/USPTO\_spool/US9873409/runat\_27032003.115416.19216/app\_query.fasta\_1.7544  
-DB=N Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US9873409 @CGN 1 1 4247 @runat\_27032003.115416.19216 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DRV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOF=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_101002.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	619	93.9	3699	24	ABK83223	Human transporter
2	472	71.6	4233	21	AAZ90198	Rat mdrlb2 (multidr
3	472	71.6	4233	22	AAF27498	Rat mdrlb2 multidr
4	469	71.2	4279	22	AAD03488	Dog P-glycoprotein
5	469	71.2	4279	22	AAD03504	Dog P-glycoprotein
6	469	71.2	4279	22	AAD03505	Dog P-glycoprotein
7	469	71.2	4279	22	AAD03506	Dog P-glycoprotein
8	468	71.0	4189	21	AAZ49334	Murine multidrug r
9	468	71.0	4189	24	ABA94367	Mouse BCRP DNA rel
10	468	71.0	4313	14	AAQ38950	Mouse multidrug re
11	467	70.9	4254	24	ABK63517	Rat sequence diffe
12	467	70.9	4369	21	AAZ52047	Rat multidrug resi
13	467	70.9	4425	21	AAZ52048	Rat multidrug resi
14	465	70.6	4186	22	AAF86127	Cynomologous monke
15	465	70.6	4195	22	AAF86128	Cynomologous monke
16	465	70.6	4317	22	AAD03489	Dog P-glycoprotein
17	462	70.1	4788	21	AAZ49335	Murine multidrug r
18	462	70.1	4788	24	ABA94368	Mouse BCRP DNA rel
19	461	70.0	3840	24	ABL91687	Human polynucleoti
20	461	70.0	3860	21	AAZ49332	Human wild-type mu
21	461	70.0	3860	21	AAZ49333	Human G185V mutant
22	461	70.0	3860	24	ABA94365	Human BCRP DNA rel
23	461	70.0	3860	24	ABA94366	Human BCRP DNA rel
24	461	70.0	3988	21	AAZ88973	Human MDR-1 DNA..
25	461	70.0	4264	19	AAV66533	Mutated human P-gl
26	461	70.0	4264	19	AAV66534	Mutated human P-gl
27	461	70.0	4349	22	AAH57442	Human intestine ce
28	461	70.0	4378	11	AAQ04522	Multidrug Resistan
29	461	70.0	4646	15	AAQ72872	Human multidrug re
30	461	70.0	4646	21	AAZ94738	Human ATP binding
31	461	70.0	4646	24	AAZ38994	Human mdri gene.
32	461	70.0	4646	24	ABL68592	Kidney cancer rela
33	461	70.0	4646	24	ABL68580	Kidney cancer rela
34	461	70.0	4669	8	AAZ70752	Sequence of human
35	461	70.0	4669	14	AAQ52726	Sequence of human
36	461	70.0	4669	19	AAV32645	Human P glycoprote
37	461	70.0	4669	24	ABK52041	cDNA encoding huma
38	461	70.0	6505	17	AAZ13394	Hybrid vector pSF-
39	461	70.0	8630	21	AAZ24041	Retroviral M4 mdr-
40	461	70.0	8630	21	AAZ24042	Retroviral vector
41	448	68.0	3912	24	ABK63653	Rat sequence diffe
42	440	66.8	3924	21	AAZ94742	Human ATP binding
43	440	66.8	3924	21	AAZ88974	Human MDR-3 DNA.
44	440	66.8	3924	24	ABN95801	Gene #2299 used to
45	427	64.8	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1  
ABK83223  
ID ABK83223 standard; cDNA; 3699 BP.  
XX  
AC ABK83223;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.  
XX  
KW Human; ss; gene; transporter and ion channel; TRICH; transport disorder;  
KW neurological disorder; muscle disorder; immunological disorder; cancer;  
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;  
KW cell proliferative disorder; cervical cancer; breast cancer;  
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;  
KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;  
KW Grave's disease; gastrointestinal disorder; Crohn's disease;  
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;  
 KW protozoal infection; helminthic infection; cardiovascular disorder;  
 KW atherosclerosis; hepatic disease.

OS Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

PR 03-NOV-2000; 2000US-245904P.

PR 09-NOV-2000; 2000US-247673P.

PR 17-NOV-2000; 2000US-249661P.

PR 20-NOV-2000; 2000US-252232P.

PR 01-DEC-2000; 2000US-250790P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

PI Walia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

PI Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

PI Thangaveilu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI: 2002-463570/49.

DR P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for  
 PT diagnosing, preventing, and treating disorders associated with an  
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular  
 PT or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
 CC a recombinant polynucleotide comprising a promoter sequence operably  
 CC linked to the TRICH polynucleotide, a cell transformed with the  
 CC recombinant polynucleotide, a transgenic organism comprising the  
 CC recombinant polynucleotide, an isolated antibody that binds specifically  
 CC to TRICH, and screening for compounds which bind to TRICH, modulate  
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.  
 CC The polypeptides are useful for diagnosing, treating, and  
 CC preventing transport, neurological, muscle, immunological disorders  
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or  
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
 CC catatonial), endocrine disorders (e.g. diabetes, Grave's disease),  
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,  
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
 CC other diseases and disorders detailed in the specification. They can also  
 CC be used in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of transporters and  
 CC ion channels. TRICH or its fragments may also be used in screening for  
 CC compounds that specifically bind to and modulate the activity of TRICH.  
 CC The polynucleotides can be used to create knock-in humanised animals or  
 CC transgenic animals to model human disease. The present sequence  
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Pred. No.: 3.68e-72 Length: 3699  
 Score: 619.00 Matches: 124  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 93.93% Indels: 0  
 DB: 24 Gaps: 0

US-09-873-409-3 (1-131) x ABK83223 (1-3699)

QY 1 MetValAspGluAenAspIleArgAlaLeuAasnValArgHisTyrArgAspHisIleGly 20  
 DB 1339 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGGCGCATTCATCGAGACCATTATTGGA 1398  
 QY 21 ValValSerGlnGluProValLeuPheGlyThrIleSerAasnAenIleLysTyrGly 40  
 DB 1399 GTGGTTAGTCAAGAGCGCTGTTTTGTCGGGACCACCATCATCAATATCAAGTATGGA 1458  
 QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAalaArgGluAalaAenAlaTyrAsp 60  
 DB 1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGACGACGAAGGAAGCAATGCGTATGAT 1518  
 QY 61 PheIleMetGluPheProAasnLysPheAasnThrLeuValGlyGluLysGlyAlaGlnMet 80  
 DB 1519 TTTATCATCGAGTTTCCTAATAAATTTAATACATTTGTTAGGGGAAAAAGAGCTCAATG 1578  
 QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAalaArgAlaLeuValArgAasnProLysIle 100  
 DB 1579 AGTGGAGGGCAGAAACAGAGGATCGCAATTCCTCGCTTGTAGTTCGAAACCCCAAGATT 1638  
 QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
 DB 1639 CTGATTTTAGATGAGGCTAGCTCTGCCCTGGATTGAGAAAGCAAGTCAGCTGTTCAAGCT 1698  
 QY 121 AlaLeuGluLys 124  
 DB 1699 GCACCTGGAGAAG 1710  
 RESULT 2  
 AAZ90198  
 ID AAZ90198 standard; DNA; 4233 BP.  
 XX AC AAZ90198;  
 XX DT 19-MAY-2000 (first entry)  
 XX DE Rat mdrlb2 (multispecific drug transporter) nucleotide sequence.  
 KW Rat; mdrlb2; multispecific drug transporter; cancer; inflammation;  
 KW cardiovascular disease; central nervous system disorder;  
 KW autoimmune disease; kidney disease; drug formulation; ss.  
 XX Rattus sp.  
 XX US6025160-A.  
 XX PD 15-FEB-2000.  
 XX PF 22-JUL-1998; 98US-0120513.  
 XX PR 22-JUL-1998; 98US-0120513.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;  
 XX WPI: 2000-181810/16.  
 XX P-PSDB; AAY78879.  
 PT Isolated polynucleotide encoding a rat mdrlb2 drug transporter  
 PT polypeptide, useful for treatment of e.g. cancer, autoimmune disease,  
 PT central nervous system disorders  
 XX Claim 7; Column 13-17, 17pp; English.  
 XX This sequence represents a polynucleotide encoding a 1275 amino acid  
 CC rat mdrlb2 multispecific drug transporter polypeptide. The mdrlb2

CC nucleotide sequence has cytostatic, antiinflammatory, cardiant,  
 CC neuroprotective, immunosuppressive and nephrotropic activity.  
 CC Underlying the functioning of the mdrlb2 polynucleotide and protein in  
 CC transgenic animal models is useful for treating and preventing diseases  
 CC such as cancer, inflammation, cardiovascular disease, central nervous  
 CC system disorders, autoimmune disease, and kidney disease. The use of the  
 CC protein in cell based, membrane based, or binding assays may enhance drug  
 CC formulation, selection of formulation excipients and compound design.  
 XX

SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores: 1.96e-52 Length: 4233  
 Pred. No.: 472.00 Matches: 91  
 Score: 472.00  
 Percent Similarity: 91.06% Conservative: 21  
 Best Local Similarity: 73.98% Mismatches: 11  
 Query Match: 71.62% Indels: 0  
 DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AA290198 (1-4233)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 DB 1376 ATCGACGGACAGACATCAGACCATCAATGTGAGGTATCTGCGGAATCATTTGGGGTG 1435  
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
 DB 1436 GTGAGTCAGAACCCCGTCTGTTTGCACACGATTCGCCGAAACATTCGCTATGCGCGA 1495  
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 DB 1496 GAAACGTCACCATGATGAGATGAGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTC 1555  
 QY 62 IleMetGluPheProAnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81  
 DB 1556 ATCATGAACCTGCCCAAAATTTAAACCCCTGTTGTTGAGAGAGGGCGCAGCTGAGT 1615  
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 DB 1616 GGGGGACAGAAACAGAGGATCGCATTTGCCGGCCCTGTTGCCCAACCCCAAGATCCTT 1675  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
 DB 1676 TTGTTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGGCGTGTTCAGGCGCT 1735  
 QY 122 LeuGlyLys 124  
 DB 1736 CTGGATAAG 1744

RESULT 3  
 AAF27498  
 ID AAF27498 standard; CDNA; 4233 BP.

XX AAF27498;  
 XX  
 XX 25-APR-2001 (first entry)  
 DT  
 XX

DE Rat mdrlb2 multidrug resistance CDNA, SEQ ID NO:1.

XX Rat; multidrug resistance; mdrlb2; multi-drug transporter family;  
 KW drug screening; pharmacokinetic analysis; oral absorption;  
 KW formulation design; bioavailability; transgenic animal; knockout animal;  
 KW inflammation; cardiovascular disease; central nervous system disorder;  
 KW cancer; autoimmune disorder; kidney disease; ss.

OS Rattus sp.  
 XX  
 XX US6169166-B1.  
 XX

XX 02-JAN-2001.

XX 29-NOV-1999; 99US-0450105.

XX

PR 22-JUL-1998; 98US-0120513.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Chenery RJ, Ellens H, Brun KA, Yue L, Peild JA;  
 XX WPI; 2001-158280/16.  
 DR P-PSDB; AAB60409.  
 XX  
 PT New rat multidrug resistance protein (mdrlb2), useful for screening or  
 PT identifying compounds that are (ant)agonists/inhibitors of the  
 PT ratmdrlb2, as well as compounds with optimal development  
 PT characteristics -  
 XX  
 PS Disclosure; Column 13-14; 13pp; English.  
 XX  
 CC The invention relates to a novel rat multidrug resistance (mdr)  
 CC protein, mdrlb2 (AAB60409). Multi-specific drug transporter family  
 CC proteins are present in cell which have a barrier function, such as  
 CC intestinal epithelial cells, brain microvessel endothelial cells,  
 CC kidney epithelial cells, and liver hepatocytes, and are also expressed  
 CC by certain cancer cells. The rat mdrlb2 protein is useful for  
 CC screening or identifying compounds that are agonists or antagonists  
 CC of mdrlb2 activity. It may also be used to establish assays to  
 CC predict oral absorption and pharmacokinetics of drugs in humans, and  
 CC thus enhance the design of formulations through the identification of  
 CC compounds with optimal development characteristics (i.e., high oral  
 CC bioavailability, UID (once a day) dosing, reduced drug interactions,  
 CC reduced variability, and reduced food effects), specifically to  
 CC avoid interactions with human mdr-1. Transgenic and knockout animals  
 CC created using DNA encoding the rat mdrlb2 may be used to gain an  
 CC insight into treating and preventing human diseases such as cancer,  
 CC inflammation, cardiovascular disease, central nervous system disorders,  
 CC autoimmune disorders and kidney disease. The present sequence represents  
 CC cDNA encoding rat mdrlb2.  
 XX

SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores: 1.96e-52 Length: 4233  
 Pred. No.: 472.00 Matches: 91  
 Score: 472.00  
 Percent Similarity: 91.06% Conservative: 21  
 Best Local Similarity: 73.98% Mismatches: 11  
 Query Match: 71.62% Indels: 0  
 DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAF27498 (1-4233)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 DB 1376 ATCGACGGACAGACATCAGACCATCAATGTGAGGTATCTGCGGAATCATTTGGGGTG 1435  
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
 DB 1436 GTGAGTCAGAACCCCGTCTGTTTGCACACGATTCGCCGAAACATTCGCTATGCGCGA 1495  
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 DB 1496 GAAACGTCACCATGATGAGATGAGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTC 1555  
 QY 62 IleMetGluPheProAnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81  
 DB 1556 ATCATGAACCTGCCCAAAATTTAAACCCCTGTTGTTGAGAGAGGGCGCAGCTGAGT 1615  
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 DB 1616 GGGGGACAGAAACAGAGGATCGCATTTGCCGGCCCTGTTGCCCAACCCCAAGATCCTT 1675  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
 DB 1676 TTGTTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGGCGTGTTCAGGCGCT 1735

```

QY 122 LeuGluLys 124
Db 1736 CTGGATAAG 1744

RESULT 4
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
AC AAD03488;
XX 13-JUN-2001 (first entry)
XX Dog P-glycoprotein (PGP) cDNA #1.
DE Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX Canis familiaris.
OS
XX Location/Qualifiers
FH Key 17..3862
FT CDS
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX WO200123540-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00303.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX Claim 3; Page 58-63; 111pp; English.
XX The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.: 5e-52 Length: 4279
Score: 469.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 22 Gaps: 0

```

---

```

US-09-873-409-3 (1-131) x AAD03488 (1-4279)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValAtcGHisTyArgAspHisIleGlyVal 21
Db 1379 ATTGATGGACAGACATTAGGACCATAAATGTAAGCACTCTCGGAAATATTACTGTGTG 1438
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyGlyArg 41
Db 1439 GTGAGTCAGAGGCTGTGTTGTCACACAGATAGCTGAAACATTCGCTATGCCGC 1498
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyArgPhe 61
Db 1499 GAAATGTCACCATGATGAGATTGAGAAAGCTGTTAAGAGAGCAATGCTATGATTTT 1558
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1559 ATCATGAAACTACCTAATAATTGACACTCTGTTGGAGAGAGGGGCCAGCTGAGT 1618
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1619 GGTGGACAGAAACAGAGAAATCGCCATTGCTCGGGCCCTGTTTCGCAACCCCAAGATTCTT 1678
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db 1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGGAAGAGTGAAGCAGTGGTTCAGTGGCC 1738
QY 122 LeuGluLys 124
Db 1739 CTGGATAAG 1747

RESULT 5
AAD03504
ID AAD03504 standard; cDNA; 4279 BP.
XX AAD03504;
XX 13-JUN-2001 (first entry)
XX Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX Canis familiaris.
XX Key Location/Qualifiers
FH Key 17..3862
FT CDS
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype A) protein"
FT allele replace (607, C)
FT /*tag= b
XX WO200123540-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00308.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX

```

PS Claim 9; Page 85-90; 111pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred

CC as multidrug transporter (MDR1) and nucleic acids encoding them.

CC The invention also includes fragments and biologically functional

CC variants of dog P-glycoprotein. PGP and their nucleic acids are

CC useful for determining the bioavailability of drugs and for

CC screening PGP inhibitors. They are useful for the diagnosis and

CC treatment of conditions characterised by PGP activity, by

CC reducing or increasing PGP activity in a cell. PGP nucleic acids

CC are used as oligonucleotide probes. Complements of PGP nucleic

CC acids are useful as antisense oligonucleotides, to induce a PGP

CC 'knockout' phenotype. They are used to prepare a non-human

CC transgenic animal, which are valuable as genetic models for

CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) allelic variant

CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump

CC exporting small molecules across the cell membrane. This enzyme

CC is a member of the ABC transporter family.

XX SQ Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:

Pred. No.: 5e-52 Length: 4279

Score: 469.00 Matches: 90

Percent Similarity: 89.43% Conservative: 20

Best Local Similarity: 73.17% Mismatches: 13

Query Match: 71.17% Indels: 0

DB: 22 Gaps: 0

US-09-873-409-3 (1-131) x AAD03504 (1-4279)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21

Db 1379 ATTGATGGACAGGACATAGGACATTAATGTAAGGCATCTTCGGGAATATCTGGTGTG 1438

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41

Db 1439 GTGAGTCAGGAGCGCTGTGTGTTTGGCCACCATGCTGAAACATTCGCTATGCGCGC 1498

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61

Db 1499 GAAATATGTCACCATGATGATGATGTAAGAGCTGTTAAGGAACCAATGCTATGATTTT 1558

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Db 1559 ATCATGAATCTACTATAATTTTGACACTCTGTTGGAGAGAGGGGCCAGCTGAGT 1618

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

Db 1619 GGTGGACAGAAACAGAGATCGCATTTGCTCGGGCCCTGTTGCAACCCCAAGATTCTT 1678

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAala 121

Db 1679 CTGCTGGATGAGGACAGCTCAGCTCTGGACACTGGAAGTGAAGCAGGTGTTTCAGTGGCC 1738

QY 122 LeuGluLys 124

Db 1739 CTGGATAAG 1747

RESULT 6

AAD03505

ID AAD03505 standard; cDNA; 4279 BP.

XX AAD03505;

XX 13-JUN-2001 (first entry)

XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.

XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;

XX MDR1; drug bioavailability; transgenic animal; genetic model; ss.

OS Canis familiaris.

XX Key Location/Qualifiers

FT CDS 17..3862

FT /\*tag= a

FT /product= "Dog P-glycoprotein (PGP) allelic variant

FT (Genotype B) protein"

FT allele replace (91, T)

FT /\*tag= b

FT allele replace (607, C)

FT /\*tag= c

XX WO200123540-A2.

PN 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26767.

XX 28-SEP-1999; 99US-0156510.

XX (GENT-) GENTEST CORP.

XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;

XX WPI; 2001-235373/24.

XX P-PSDB; AAE00309.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful

PT for determining the bioavailability of drugs and for screening for dog

PT PGP inhibitors -

XX Claim 9; Page 93-99; 111pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred

CC as multidrug transporter (MDR1) and nucleic acids encoding them.

CC The invention also includes fragments and biologically functional

CC variants of dog P-glycoprotein. PGP and their nucleic acids are

CC useful for determining the bioavailability of drugs and for

CC screening PGP inhibitors. They are useful for the diagnosis and

CC treatment of conditions characterised by PGP activity, by

CC reducing or increasing PGP activity in a cell. PGP nucleic acids

CC are used as oligonucleotide probes. Complements of PGP nucleic

CC acids are useful as antisense oligonucleotides, to induce a PGP

CC 'knockout' phenotype. They are used to prepare a non-human

CC transgenic animal, which are valuable as genetic models for

CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) allelic variant

CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump

CC exporting small molecules across the cell membrane. This enzyme

CC is a member of the ABC transporter family.

XX SQ Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:

Pred. No.: 5e-52 Length: 4279

Score: 469.00 Matches: 90

Percent Similarity: 89.43% Conservative: 20

Best Local Similarity: 73.17% Mismatches: 13

Query Match: 71.17% Indels: 0

DB: 22 Gaps: 0

US-09-873-409-3 (1-131) x AAD03505 (1-4279)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21

Db 1379 ATTGATGGACAGGACATAGGACATTAATGTAAGGCATCTTCGGGAATATCTGGTGTG 1438

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41

Db 1439 GTGAGTCAGGAGCGCTGTGTGTTTGGCCACCATGCTGAAACATTCGCTATGCGCGC 1498

QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61

Db 1499 GAAATATGTCACCATGATGATGATGTAAGAGCTGTTAAGGAACCAATGCTATGATTTT 1558

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Db 1559 ATCATGAATCTACTATAATTTTGACACTCTGTTGGAGAGAGGGGCCAGCTGAGT 1618

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

Db 1619 GGTGGACAGAAACAGAGATCGCATTTGCTCGGGCCCTGTTGCAACCCCAAGATTCTT 1678

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAala 121

Db 1679 CTGCTGGATGAGGACAGCTCAGCTCTGGACACTGGAAGTGAAGCAGGTGTTTCAGTGGCC 1738

QY 122 LeuGluLys 124

Db 1739 CTGGATAAG 1747

RESULT 6

AAD03505

ID AAD03505 standard; cDNA; 4279 BP.

XX AAD03505;

XX 13-JUN-2001 (first entry)

XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.

XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;

XX MDR1; drug bioavailability; transgenic animal; genetic model; ss.



Db 1499 GAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTTT 1558  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1559 ATCATGAACTACCTATATAATTTGACACTCTGTTGGAGAGAGAGGGCCAGCTGAGT 1618  
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1619 GGTGGACAGAAACAGAGAAATCGCATTTGCTCGGGCCCTGGTTGCAACCCCAAGATTCTT 1678  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 1679 CTGCTGGATGAGGCAACGTGAGTCTTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGCC 1738  
 QY 122 LeuGluLys 124  
 Db 1739 CTGGATAAG 1747  
 RESULT 7  
 AAD03506  
 ID AAD03506 standard; cDNA; 4279 BP.  
 XX  
 AC AAD03506;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.  
 XX  
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KM MDRI; drug bioavailability; transgenic animal; genetic model; ss.  
 XX  
 OS Canis familiaris.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 17..3862  
 FT /\*tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 (Genotype D) protein"  
 FT allele replace (91, T)  
 FT /\*tag= b  
 FT allele replace (607, C)  
 FT /\*tag= c  
 FT allele replace (1001, T)  
 FT /\*tag= c  
 FT allele replace (3458, A)  
 FT /\*tag= c  
 XX  
 PN WO200123540-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PP 28-SEP-2000; 2000WO-US26767.  
 XX  
 PR 28-SEP-1999; 99US-0156510.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
 XX  
 DR WPI; 2001-235373/24.  
 DR P-PSDB; AAE00310.  
 XX  
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 for determining the bioavailability of drugs and for screening for dog  
 PGP inhibitors -  
 XX  
 PS Claim 9; Page 102-107; Illipp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 as multidrug transporter (MDRI) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 variants of dog P-glycoprotein. PGP and their nucleic acids are  
 useful for determining the bioavailability of drugs and for

CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 XX  
 SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5e-52 Length: 4279  
 Score: 489.00 Matches: 90  
 Percent Similarity: 89.43% Conservative: 20  
 Best Local Similarity: 73.17% Mismatches: 13  
 Query Match: 71.17% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-873-409-3 (1-131) x AAD03506 (1-4279)  
 QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1379 ATTGATGGACAGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTAAGTGTG 1438  
 QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1439 GTGAGTCAGAGCCTGTGTTTGGCCACCATGATAGCTGAAACCATTCGTTATGCGCCG 1498  
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaAArgGluAlaAsnAlaTyrAspPhe 61  
 Db 1499 GAAATGTTTACCATGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTT 1558  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1559 ATCATGAACTACCTATATAATTTGACACTCTGTTGGAGAGAGAGGGCCAGCTGAGT 1618  
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1619 GGTGGACAGAAACAGAGAAATCGCCATTGCTCGGGCCCTGGTTGCAACCCCAAGATTCTT 1678  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 1679 CTGCTGGATGAGGCAACGTGAGTCTTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGCC 1738  
 QY 122 LeuGluLys 124  
 Db 1739 CTGGATAAG 1747  
 RESULT 8  
 AAZ49334  
 ID AAZ49334 standard; cDNA; 4189 BP.  
 XX  
 AC AAZ49334;  
 XX  
 DT 14-MAR-2000 (first entry)  
 XX  
 DE Murine multidrug resistance-1 (MDR-1) cDNA.  
 XX  
 KW Multidrug resistance; MDR-1; P-glycoprotein;  
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassaemia;  
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
 KW cytokine; ds.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers

FT CDS 1..3831  
 FT /\*tag= a  
 FT /product= "Murine MDR-1 protein"  
 XX  
 PN W09961589-A2.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 27-MAY-1999; 99WO-US11825.  
 XX  
 PP 28-MAY-1998; 98US-0086988.  
 XX  
 PR (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA  
 PI Sorrentino B, Bunting K;  
 DR WPI: 2000-072615/06.  
 DR P-PSDB; AAY58188.  
 XX  
 PT Ex vivo expansion of hematopoietic stem cells transduced with a  
 PT sequence encoding human multidrug resistance-1, used for bone marrow  
 PT transplantation -  
 PS Disclosure; Page 90-93; 113pp; English.  
 XX  
 CC This sequence represents cDNA encoding murine multidrug  
 CC resistance protein MDR-1. MDR-1 is a transmembrane  
 CC efflux pump, responsible for the export of drugs from cells,  
 CC particularly cancer cells. The invention relates to transducing  
 CC hematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified hematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC hematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from hematopoietic stem cells, e.g., thalassemia,  
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in hematopoietic stem cells.  
 CC Hematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC hematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.  
 XX  
 SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:  
 Pred. No.: 6..6e-52 Length: 4189  
 Score: 468.00 Matches: 90  
 Percent Similarity: 91.06% Conservative: 22  
 Best Local Similarity: 73.17% Mismatches: 11  
 Query Match: 71.02% Indels: 0  
 DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAZ49334 (1-4189)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 DB 1354 ATCCGACGGACAGACATCAGAACCAATCAATGTGAGGTATCTGAGGGGAGATCAATGGTGTG 1413  
 QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 DB 1414 GTGAGTCAGAACTGTGCTGTTTGGCCACACGATCCCGAGAACATTCCTGATGCCGA 1473  
 QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 DB 1474 GAAGATGTCAACATGGATGAGATTGAGAAAGCTCTCAAGGAGCAATGCTATGACTTC 1533  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 DB 1534 ATCATGAACCTGCCCACTTTCACACCTCTGTTGTTGAGAGAGGGGCGCAGCTGAGT 1593

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 DB 1594 GGGGACAGAAACAGAGATCCCATTCGCGGCCCTGTCGCAATCCCAAGATCCTT 1653  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 121  
 DB 1654 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGACGGCGCA 1713  
 QY 122 LeuGluLys 124  
 DB 1714 CTGGATAAG 1722  
 RESULT 9  
 ABA94367  
 ID ABA94367 standard; DNA; 4189 BP.  
 XX  
 AC ABA94367;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Mouse BCRP DNA related seq Id No. 5.  
 XX  
 KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;  
 KW hepatotropic; neurodegenerative; cytosolic; antianemic; muscular; BCRP;  
 KW cardiant; gene therapy; ds.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..3831  
 FT /\*tag= a  
 XX  
 PN W0200192877-A2.  
 PD 06-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US17459.  
 XX  
 PR 31-MAY-2000; 2000US-0584586.  
 PR 29-MAY-2001; 2001US-0866866.  
 XX  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA  
 PI Sorrentino B, Schuetz J;  
 DR WPI: 2002-114368/15.  
 DR P-PSDB; ABB07268.  
 XX  
 PT Identifying a stem cell, for treating e.g., muscular dystrophy,  
 PT myocardial infarction, Parkinson's disease, or neurodegenerative  
 PT disorders, comprises detecting the expression of an ATP transport  
 PT protein (BCRP) by a cell -  
 XX  
 PS Disclosure; Page 64-66; 87pp; English.  
 XX  
 CC The invention provides a method of identifying and/or isolating a stem  
 CC cell that involves detecting the expression of an ATP transport protein  
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample  
 CC comprising stem cells. The isolated stem cells may be used in the  
 CC treatment of diseases such as muscular dystrophy, degenerative liver  
 CC disorder, myocardial infarction, Parkinson's disease, degenerative  
 CC disorders of the brain, and for tissue regeneration or replacement.  
 CC Hematopoietic cells can be used in bone marrow transplants (e.g., for  
 CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
 CC also be used as cell targets in gene therapy protocols. The present  
 CC sequence represents a sequence related to the BCRP for which no relevant  
 CC information has been provided in the specification.  
 XX  
 SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:

```

Pred. No.: 6.6e-52 Length: 4189
Score: 468.00 Matches: 90
Percent Similarity: 91.06% Conservative: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 71.02% Indels: 0
DB: 24 Gaps: 0

US-09-873-409-3 (1-131) x ABA94367 (1-4189)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1354 ATCCAGCGGACAAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTG 1413
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1414 GTGAGTCAGGAACTGTGCTGTTTGCACACAGATCCCGGAGACATTCGCTATGGCCGA 1473
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1474 GAAGATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCTATGACTTC 1533
QY 62 IleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMetSer 81
DB 1534 ATCATGAAACTGCCCACTTGTGACACCTGTTGTGAGAGAGGGGGCGAGCTGAGT 1593
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1594 GGGGACAGAAACAGAGATCGCATTCGCGGCGCTGTGTCGCCAATCCCAAGATCCTT 1653
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1654 TTGTTGGACGAGGCGCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTCAGGCGCGCA 1713
QY 122 LeuGluLys 124
DB 1714 CTGGATAAG 1722

RESULT 10
AAQ38950
ID AAQ38950 standard; DNA; 4313 BP.
AC AAQ38950;
XX
XX
XX 28-JUL-1993 (first entry)
DE Mouse multidrug resistance sequence.
XX
XX mdr gene; Lambda DR11 clone; ss.
XX
XX Mus musculus.
FH Key Location/Qualifiers
FT CDS 110..3940
FT FT /*tag= a
FT FT /phenotype= multidrug_resistance
XX
XX
XX US5198344-A.
XX
XX 30-MAR-1993.
XX
XX 15-JUL-1986; 86US-0885951.
XX
XX 15-JUL-1986; 86US-0885951.
XX
XX 06-FEB-1991; 91US-0652311.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Croop JM, Gros P, Housman DE;
XX
XX WPT; 1993-126077/15.
XX
XX P-PSDB; AAR35199.
XX
XX DNA sequence which confers multi-drug resistance on sensitive

```

```

PT mammalian cells - used to preserve bone marrow cells during
PT chemotherapy to prevent infection
XX
XX Claim 1; Fig 8; 22pp; English.
XX
XX A cDNA library was constructed from mRNA which had been isolated
XX from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
XX complementary to the mRNA species encoded by 2 related but distinct
XX mdr genes were isolated and cloned. One of the cDNA clones (lambda
XX DR11) is a full-length cDNA clone for one member of the mdr gene
XX family. When incorporated into prokaryotic expression vector pDREX4
XX (which allows high levels of transcription of the cDNA when
XX introduced into mammalian cells), the clone was shown to confer the
XX multidrug resistance phenotype upon transfection into drug-sensitive
XX mammalian cells.
XX
XX Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 6.87e-52 Length: 4313
Score: 468.00 Matches: 90
Percent Similarity: 91.06% Conservative: 22
Best Local Similarity: 73.17% Mismatches: 11
Query Match: 71.02% Indels: 0
DB: 14 Gaps: 0

US-09-873-409-3 (1-131) x AAQ38950 (1-4313)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1463 ATCCAGCGGACAAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTG 1522
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1523 GTGAGTCAGGAACTGTGCTGTTTGCACACAGATCCCGGAGACATTCGCTATGGCCGA 1582
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1583 GAAGATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCTATGACTTC 1642
QY 62 IleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMetSer 81
DB 1643 ATCATGAAACTGCCCACTTGTGACACCTGTTGTGAGAGAGGGGGCGAGCTGAGT 1702
QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1703 GGGGACAGAAACAGAGATCGCATTCGCGGCGCTGTGTCGCCAATCCCAAGATCCTT 1762
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1763 TTGTTGGACGAGGCGCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTCAGGCGCGCA 1822
QY 122 LeuGluLys 124
DB 1823 CTGGATAAG 1831

RESULT 11
ABK63517
ID ABK63517 standard; cDNA; 4254 BP.
XX
XX AC ABK63517;
XX
XX 18-JUN-2002 (first entry)
XX
XX Rat sequence differentially expressed in response to a hepatotoxin #1424.
XX
XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX
XX differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
XX
XX OS
XX
XX WO200210453-A2.
XX
XX

```

PD 07-FEB-2002.  
 XX  
 PF 30-JUL-2001; 2001WO-US23872.  
 XX  
 PR 31-JUL-2000; 2000US-222040P.  
 PR 02-NOV-2000; 2000US-244880P.  
 PR 11-MAY-2001; 2001US-290029P.  
 PR 15-MAY-2001; 2001US-290645P.  
 PR 22-MAY-2001; 2001US-292336P.  
 PR 06-JUN-2001; 2001US-295798P.  
 PR 13-JUN-2001; 2001US-297457P.  
 PR 19-JUN-2001; 2001US-298884P.  
 PR 09-JUL-2001; 2001US-303459P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
 XX  
 DR WPI; 2002-241625/29.  
 XX  
 PT Predicting toxic effects of compounds or the progression of these toxic  
 PT effects by determining the changes in gene expression in tissues or  
 PT cells exposed to the toxin and comparing these to gene expression in  
 PT unexposed tissues or cells -  
 XX  
 PS Claim 1; Seq ID No 1424; 239pp; English.  
 XX  
 CC The invention relates to methods for predicting toxic effects of  
 CC compounds or the progression of these toxic effects by determining the  
 CC global changes in gene expression in tissues or cells exposed to the  
 CC toxin and comparing these to gene expression in unexposed tissues or  
 CC cells. Also included are methods of predicting at least one toxic  
 CC effect of a compound or progression of a toxic effect, preferably the  
 CC hepatotoxicity of a compound, comprising detecting the level of  
 CC expression in a tissue or cell sample exposed to the compound of two or  
 CC more genes listed in the specification, where differential expression of  
 CC the genes is indicative of at least one toxic effect or progression.  
 CC The method can also be used to identify an agent which modulates the  
 CC toxic response and predict cellular pathways that a compound modulates  
 CC in a cell. The methods utilise a set of at least two probes (on a solid  
 CC support in kit form) where each of the probes comprises a sequence that  
 CC specifically hybridises to a gene listed in the specification, a computer  
 CC system comprising a database containing information identifying the  
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
 CC set of genes comprising at least two genes listed in the specification,  
 CC and a user interface to view the information used to present information,  
 CC identifying the expression level in a tissue or cell of at least one gene  
 CC listed in the specification. The method is useful for elucidating global  
 CC changes in gene expression and for identifying toxicity markers in  
 CC tissues or cell exposed to a known toxin. The genes may be used as  
 CC toxicity markers in drug screening and toxicity assays. The genes and  
 CC gene expression information may be used as diagnostic markers for the  
 CC prediction or identification of the physiological state of tissue or cell  
 CC sample that has been exposed to a compound or agent. Hepatotoxicity  
 CC is characterised by centrilobular necrosis and steatosis. The present  
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
 CC which is differentially expressed in response to a hepatotoxic agent.  
 XX  
 SQ Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 9.17e-52 Length: 4254  
 Score: 467.00 Matches: 90  
 Percent Similarity: 91.06% Conservative: 22  
 Best Local Similarity: 73.17% Mismatches: 11  
 Query Match: 70.86% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-873-409-3 (1-131) x ABR63517 (1-4254)  
 QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 Db 1456 ATTGACGGAGCAGACATCAGGACCATCAATGTGAGGTATCTCGGGAATCATTTGGGGTG 1515

QY 22 ValSerGlnGluProValIleuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 Db 1516 GTGAGTCAGGAACCGCTGCTGTTTCCACCGATTCGCGAAACATTGCTATGCGCGA 1575  
 QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 Db 1576 GAAACGTCACCATGATGAGATGAGAAAGCTGTCAAGGAAGCAATGCCATGACTTC 1635  
 QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
 Db 1636 ATCATGAAACTGCCCCACAAATTTGACACCCCTGGTGTGAGAGAGGGCGCAGCTGACT 1695  
 QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 Db 1696 GGGGACAGAAACAGAGGATGCCATTGCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1755  
 QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 Db 1756 TTGTTGGGATAGCCACGTCAGCCTTGACACACAGAAAGCAAGCCGTGTTCAGGCCGCT 1815  
 QY 122 LeuGluLys 124  
 Db 1816 CTGGATAAG 1824  
 RESULT 12  
 AAZ52047  
 ID AAZ52047 standard; cDNA; 4369 BP.  
 AC AAZ52047;  
 XX  
 DT 18-JUL-2000 (first entry)  
 XX  
 DE Rat multidrug resistance protein 1a cDNA.  
 XX  
 KW Multidrug resistance protein 1a; mdria; multi-specific drug transporter;  
 KW drug formulation; formulation excipient; compound design; inflammation;  
 KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;  
 KW central nervous system disorder; auto-immune disease; kidney disease; ss.  
 XX  
 OS Rattus rattus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 352..4170  
 FT /\*tag= a  
 FT /product= "Rat multidrug resistance protein"  
 XX  
 PN WO200015650-A1.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 10-SEP-1999; 99WO-US20770.  
 XX  
 PR 17-SEP-1998; 98US-0156800.  
 PR 09-DEC-1998; 98US-0208809.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;  
 XX  
 DR WPI; 2000-271372/23.  
 DR P-PSDB; AAY70596.  
 XX  
 PT Isolated rat mdria polynucleotides and polypeptides, useful in assays  
 PT to provide information on drug formulation, selection of formulation  
 PT excipients and compound design -  
 XX  
 PS Claim 2; Page 23-24; 33pp; English.  
 CC  
 CC The present cDNA sequence encodes rat multidrug resistance  
 CC protein 1a (mdria). This protein is a member of the multi-specific drug  
 CC transporters family. Mdria is used in assays to provide information on  
 CC drug formulation, selection of formulation excipients and compound

CC design. They are used in cell based, membrane based, binding or other  
CC assays to provide information that may enhance drug formulation. This  
CC invention further relates to the generation of in vivo and in vitro  
CC comparison data to predict oral absorption and pharmacokinetics. This  
CC enables the selection of drugs with optimal pharmacokinetics, i.e. good  
CC oral bioavailability, brain penetration, plasma half life, and minimum  
CC drug interaction. Transgenic and knock-out animals created using rat  
CC mdrla provides an insight into treating and preventing human diseases  
CC including cancer, inflammation, cardiovascular disease, central nervous  
CC system disorders, auto-immune and kidney disease.

XX SQ Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:  
Pred. No.: 9,51e-52 Length: 4369  
Score: 467.00 Matches: 90  
Percent Similarity: 91.06% Conservative: 22  
Best Local Similarity: 73.17% Mismatches: 11  
Query Match: 70.86% Indels: 0  
DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAZ52047 (1-4369)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
DB 1684 ATCGACGGACAGACATCAGACCATCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1743  
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
DB 1744 GTGAGTCAGAACCCCGTGTGTTGCCACCAATTCGCCGAAACATTCGTATGCGCGA 1803  
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
DB 1804 GAAACGTCACCATGATGATAGAGAAAGCTGTCAAGGAAGCCATGCTATGATTTTC 1863  
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
DB 1864 ATCATGAACTGCCCAAAATTTGACCCCTGTTGGTGAGAGAGGGGCGAGCTGAGT 1923  
QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101  
DB 1924 GGGGACAGAAACAGAGGATCGCATTCGCCGGCCCTGTCGCAACCCCAAGATCCTT 1983  
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
DB 1984 TTGTTGGATGAGGCCACGTGAGCCTTGACACACAGAAAGCGGCTGTTAGGCGCGCT 2043

QY 122 LeuGluLys 124

DB 2044 CTGGATAG 2052

RESULT 13

AAZ52048

ID AAZ52048 standard; cDNA; 4425 BP.

XX AC AAZ52048;

XX 18-JUL-2000 (first entry)

DE DE Rat multidrug resistance protein 1a cDNA derived from EST sequences.

XX Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;  
KW drug formulation; formulation excipient; compound design; inflammation;  
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;  
KW central nervous system disorder; auto-immune disease; kidney disease;  
KW EST; expressed sequence tag; ss.

XX Rattus rattus.

XX Key Location/Qualifiers

FT CDS 352..4170

FT /\*tag= a

FT /product= "Rat multidrug resistance protein"

XX WO200015650-A1.  
XX 23-MAR-2000.  
XX 10-SEP-1999; 99WO-US20770.  
XX 17-SEP-1998; 98US-0156800.  
XX 09-DEC-1998; 98US-0208809.  
XX (SMIK ) SMITHLINE BEECHAM CORP.  
XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;  
XX WPI: 2000-271372/23.  
XX P-PSDB; AAY70597.  
XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays  
XX to provide information on drug formulation, selection of formulation  
XX excipients and compound design -  
XX Claim 11; Page 27-29; 33pp; English.

XX The present cDNA sequence encodes rat multidrug resistance protein 1a  
CC (mdrla). This cDNA is derived from EST (expressed sequence tag)  
CC sequences. Mdr1a is used in assays to provide information on  
CC drug formulation, selection of formulation excipients and compound  
CC design. They are used in cell based, membrane based, binding or other  
CC assays to provide information that may enhance drug formulation. This  
CC invention further relates to the generation of in vivo and in vitro  
CC comparison data to predict oral absorption and pharmacokinetics. This  
CC enables the selection of drugs with optimal pharmacokinetics, i.e. good  
CC oral bioavailability, brain penetration, plasma half life, and minimum  
CC drug interaction. Transgenic and knock-out animals created using rat  
CC mdrla provides an insight into treating and preventing human diseases  
CC including cancer, inflammation, cardiovascular disease, central nervous  
CC system disorders, auto-immune and kidney disease.

XX SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:

Pred. No.: 9,68e-52 Length: 4425  
Score: 467.00 Matches: 90  
Percent Similarity: 91.06% Conservative: 22  
Best Local Similarity: 73.17% Mismatches: 11  
Query Match: 70.86% Indels: 0  
DB: 21 Gaps: 0

US-09-873-409-3 (1-131) x AAZ52048 (1-4425)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
DB 1684 ATCGACGGACAGACATCAGGACCATCAATGTGAGGTATCTGCGGGAATCATTTGGGGTG 1743  
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41  
DB 1744 GTGAGTCAGAACCCCGTGTGTTGCCACCAATTCGCCGAAACATTCGTATGCGCGA 1803  
QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
DB 1804 GAAACGTCACCATGATGATAGAGAAAGCTGTCAAGGAAGCCATGCTATGATTTTC 1863  
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81  
DB 1864 ATCATGAACTGCCCAAAATTTGACCCCTGTTGGTGAGAGAGGGGCGAGCTGAGT 1923  
QY 82 GlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeu 101  
DB 1924 GGGGACAGAAACAGAGGATCGCATTCGCCGGCCCTGTCGCAACCCCAAGATCCTT 1983  
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaA 121  
DB 1984 TTGTTGGATGAGGCCACGTGAGCCTTGACACACAGAAAGCGGCTGTTAGGCGCGCT 2043

QY 122 LeuGluLys 124  
 DB 2044 CTGGATAAG 2052

RESULT 14  
 ID AAF86127 standard; cDNA; 4186 BP.  
 XX  
 AC AAF86127;  
 DT 25-JUN-2001 (first entry)  
 XX  
 DE Cynomologous monkey P-glycoprotein cDNA.  
 XX  
 KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;  
 KW efflux pump; ss.  
 XX  
 OS Macaca fascicularis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 100..3942  
 FT /tag= a  
 FT /product= "PGP"  
 FT /note= "P-glycoprotein"  
 XX  
 PN WO200123565-A1.  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26592.  
 PR 28-SEP-1999; 99US-0156921.  
 PR 12-OCT-1999; 99US-0158818.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 XX Stocker PJ, Steimel-Crespi DT, Crespi CL;  
 DR WPI; 2001-316136/33.  
 DR P-PSDB; AAB81064.  
 XX  
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 cell -  
 XX  
 PS Example 1; Page 51-57; 84pp; English.  
 CC This invention relates to a polynucleotide sequence encoding a  
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention  
 CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents cDNA encoding  
 CC cynomologous monkey P-glycoprotein.  
 XX  
 SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.66e-51 Length: 4186  
 Score: 465.00 Matches: 91  
 Percent Similarity: 88.62% Conservative: 18  
 Best Local Similarity: 73.98% Mismatches: 14  
 Query Match: 70.56% Indels: 0  
 DB: 22 Gaps: 0

us-09-873-409-3 (1-131) x AAF86127 (1-4186)

QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21  
 DB 1456 GTTGATGGACAGGATATTAGGACCATAAACGTAAGTTTCTACGGGAATCATCGGTGTG 1515

QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41  
 DB 1516 GTGAGTCAGGAACCTGTATTGTTGCCACCGATAGCTGAAACACATTCCGCTATGCTGT 1575

QY 42 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61  
 DB 1576 GAAGATGTCACCATGATGAGATTGAGAAAGCTGTCAGGAAGCCATGCCCTATGACTTT 1635

QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 81  
 DB 1636 ATCATGAACTCCCTCAGAAATTTGACACCTGTTGGAGAGAGAGAGGGCCCGAGCTGAGT 1695

QY 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101  
 DB 1696 GGTGGCGCAGAACGACAGGATGCCATTGCATGCCCTGTTGCGAACCCCAAGATCCTC 1755

QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121  
 DB 1756 CTGCTGGACGAGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAAGGTGCT 1815

QY 122 LeuGluLys 124  
 DB 1816 CTGGATAAG 1824

RESULT 15  
 ID AAF86128 standard; cDNA; 4195 BP.  
 XX  
 AC AAF86128;  
 XX  
 DT 25-JUN-2001 (first entry)  
 XX  
 DE Cynomologous monkey P-glycoprotein cDNA variant 1.  
 XX  
 KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;  
 KW efflux pump; ss.  
 XX  
 OS Macaca fascicularis.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 100..3951  
 FT /tag= a  
 FT /product= "PGP"  
 FT /note= "P-glycoprotein"  
 FT /tag= b  
 FT /note= "Insertion of 9 nucleotides relative to PGP  
 allelic variant AAF86127"  
 XX  
 PN WO200123565-A1.  
 PD 05-APR-2001.  
 PF 28-SEP-2000; 2000WO-US26592.  
 PR 28-SEP-1999; 99US-0156921.  
 PR 12-OCT-1999; 99US-0158818.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 XX Stocker PJ, Steimel-Crespi DT, Crespi CL;  
 DR WPI; 2001-316136/33.  
 DR P-PSDB; AAB81065.  
 XX  
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in

```
PT cell -
XX
PS
XX
XX
XX
CC This invention relates to a polynucleotide sequence encoding a
CC cynomolgous monkey P-glycoprotein (PGP), and an allelic variant of the
CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC exporting small molecules across the cell membrane. The invention
CC includes a cynomolgous monkey (Macaca fascicularis) PGP coding sequence
CC and protein, and also that of an allelic variant. The PGP polynucleotide
CC sequence is useful for increasing PGP transporter activity in a cell.
CC Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC activity in a mammalian cell. They may also be used for increasing the
CC bioavailability of a drug. The present sequence represents cDNA encoding
CC cynomolgous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC insert compared to the PGP allelic variant given in AAF86127.
XX
SQ Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.: 1.66e-51 Length: 4195
Score: 465.00 Matches: 91
Percent Similarity: 88.62% Conservative: 18
Best Local Similarity: 73.98% Mismatches: 14
Query Match: 70.56% Indels: 0
DB: 22 Gaps: 0

US-09-873-409-3 (1-131) x AAF86128 (1-4195)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
DB 1465 GTTGATGCAGACGATATTAGACCATAAAGCTAAGGTTTCTACGGGAATCATCGGTGTG 1524
QY 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
DB 1525 GTGAGTCAGGAACCTGTATTGTTTGGCCACCACGATAGCTGAAAAACATTGCTATGGTCGT 1584
QY 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
DB 1585 GAAGATGTCACCATGAGTATGAGATTTGACAACTTGCACCTGTTGGACAGAGGGGCCAGCTGAGT 1644
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
DB 1645 ATCATGAACCTGCTCAGAAATTGACCCCTGTTGGACAGAGGGGCCAGCTGAGT 1704
QY 82 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
DB 1705 GGTGGGAGAGAGCAGAGGATCGCCATTGCACGTGCCCTGTTGGCAACCCCAAGATCCTC 1764
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
DB 1765 CTGCTGACGAGGCCACGTGAGCCTTGGACACAGAAAGTGAAGCGATGGTTCAGGTGGCT 1824
QY 122 LeuGluLys 124
DB 1825 CTGGATAAG 1833
```

Search completed: March 30, 2003, 03:08:18  
Job time : 142.031 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 1770.98 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-3  
Perfect score: 659  
Sequence: 1 MYDENDIRALNVRHYRHIG.....SEKSAVQAALKEKTPRYSF 131

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DRV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003.115418.19228/app\_query.fasta.1.7544  
-DB=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1.1 54887 @runat\_27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.cm.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	659	100.0	1175	6	AX339029	Sequence	AX339029 Sequence
2	659	100.0	1940	6	AX339034	Sequence	AX339034 Sequence
3	659	100.0	2021	6	AX339033	Sequence	AX339033 Sequence
4	619	93.9	2856	6	AX339028	Sequence	AX339028 Sequence
5	619	93.9	3177	6	AX339030	Sequence	AX339030 Sequence
6	619	93.9	3621	6	AX339032	Sequence	AX339032 Sequence
7	619	93.9	3899	6	AX478104	Sequence	AX478104 Sequence
8	619	93.9	3702	6	AX339031	Sequence	AX339031 Sequence
9	474	71.6	3489	4	AB029153	Felis cat	AB029153 Felis cat
10	472	71.6	3888	10	AY082609	Rattus no	AY082609 Rattus no
11	472	71.6	4233	6	AR123273	Sequence	AR123273 Sequence
12	469	71.2	2752	5	AF099732	Fundulus	AF099732 Fundulus
13	469	71.2	3334	4	AF419568	Canis fam	AF419568 Canis fam
14	469	71.2	4018	5	GGA9799	Gallus ga	AJ009799 Gallus ga
15	469	71.2	4045	12	AF269224	Synthetic	AF269224 Synthetic
16	469	71.2	4279	6	AX105057	Sequence	AX105057 Sequence
17	469	71.2	4279	6	AX105078	Sequence	AX105078 Sequence
18	469	71.2	4279	6	AX105080	Sequence	AX105080 Sequence
19	469	71.2	4279	6	AX105082	Sequence	AX105082 Sequence
20	468	71.0	4189	6	AX322791	Sequence	AX322791 Sequence
21	468	71.0	4298	10	MUSMDR	Mouse multi	M14757 Mouse multi
22	468	71.0	4356	10	MUSMDRAA	Mouse multi	M30697 Mouse multi
23	467	70.9	3858	4	OAU78609	Ovis aries	U78609 Ovis aries
24	467	70.9	3987	10	CRUPGP1185	Chinese ham	M5254 Chinese ham
25	467	70.9	4254	6	AX401748	Sequence	AX401748 Sequence
26	467	70.9	4254	10	RATMDRM	Rat mdr mrn	M81855 Rat mdr mrn
27	467	70.9	4280	10	CRUPGPII	C.griseus p	M60041 C.griseus p
28	467	70.9	4296	10	CRUPGPI	C.griseus p	M60040 C.griseus p
29	467	70.9	4304	10	CRUPGP1165	Chinese ham	M59253 Chinese ham
30	467	70.9	4323	10	AF286167	Rattus no	AF286167 Rattus no
31	467	70.9	4927	10	AF257746	Rattus no	AF257746 Rattus no
32	465	70.6	2005	4	AB066299	Canis fam	AB066299 Canis fam
33	465	70.6	4186	6	AX108654	Sequence	AX108654 Sequence
34	465	70.6	4195	6	AX108656	Sequence	AX108656 Sequence
35	465	70.6	4317	4	AF045016	Canis fam	AF045016 Canis fam
36	465	70.6	4317	6	AX105059	Sequence	AX105059 Sequence
37	465	70.6	4390	5	XLU17608	Xenopus lae	U17608 Xenopus lae
38	462	70.1	3682	10	MUSPGLY1A	Mouse phosp	M24417 Mouse phosp
39	462	70.1	4788	6	AX322793	Sequence	AX322793 Sequence
40	462	70.1	4924	10	MUSMDR1A	Mouse p-gly	M33581 Mouse p-gly
41	461	70.0	3840	6	AX481416	Sequence	AX481416 Sequence
42	461	70.0	3860	6	AX322787	Sequence	AX322787 Sequence
43	461	70.0	3860	6	AX322789	Sequence	AX322789 Sequence
44	461	70.0	3988	6	AX024454	Sequence	AX024454 Sequence
45	461	70.0	4192	9	AF016535	Homo sapi	AF016535 Homo sapi

ALIGNMENTS

RESULT 1

AX339029  
LOCUS AX339029 1175 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 11 from Patent WO0194400.  
ACCESSION AX339029  
VERSION AX339029.1 GI:18129121  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 11 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source 1. .1175  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 337 a 238 c 258 g 342 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.24e-60 Length: 1175  
Score: 659.00 Matches: 131  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-3 (1-131) x AX339029 (1-1175)  
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
Db 418 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGCGCATTTATCGAGACCATATTGGA 477  
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
Db 478 GTGGTTAGTCAAGAGCCCTGTTTGTTCGGGACCACCATCATGTAACAATATCAAGTATGGA 537  
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
Db 538 CGAGATGATGTGACTGTAGTGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATCGCTATGAT 597  
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMet 80  
Db 598 TTTATCATGGAGTTTCCCTAAATAATTTAAATACATTGGTAGGGGAAAAAGGAGCTCAAAATG 657  
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
Db 658 AGTGGAGGGCAGAAACACAGAGATCGCAATTCCTGTGCTTAGTTCGAAACCCCAAGATT 717  
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerAlaValGlnAla 120  
Db 718 CTGATTTTATAGTAGGCTACGCTCCCTCGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 777  
Qy 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131  
Db 778 GCACCTGGAGGAAGGATACCCCGAGGTATTCATTT 810  
RESULT 2  
AX339034  
LOCUS AX339034 1940 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 16 from Patent WO0194400.  
ACCESSION AX339034  
VERSION AX339034.1 GI:18129126  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE

AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 16 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source 1. .1940  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 530 a 394 c 448 g 568 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7.66e-60 Length: 1940  
Score: 659.00 Matches: 131  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-3 (1-131) x AX339034 (1-1940)  
Qy 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20  
Db 1183 ATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGCGCATTTATCGAGACCATATTGGA 1242  
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40  
Db 1243 GTGGTTAGTCAAGAGCTGTTTGTTCGGGACCACCATCATGTAACAATATCAAGTATGGA 1302  
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60  
Db 1303 CGAGATGATGTGACTGTAGTGAAGAGATGAGAGAGCAGCAAGGGAAGCAAAATCGTATGAT 1362  
Qy 61 PheIleMetGluPheProAsnLysPheAsnThrIleuValGlyGluLysGlyAlaGlnMet 80  
Db 1363 TTTATCATGGAGTTTCCCTAAATAATTTAAATACATTGGTAGGGGAAAAAGGAGCTCAAAATG 1422  
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100  
Db 1423 AGTGGAGGGCAGAAACACAGAGATCGCAATTCCTGTGCTTAGTTCGAAACCCCAAGATT 1482  
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
Db 1483 CTGATTTTATAGTAGGCTACGCTCCCTCGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542  
Qy 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131  
Db 1543 GCACCTGGAGGAAGGATACCCCGAGGTATTCATTT 1575  
RESULT 3  
AX339033  
LOCUS AX339033 2021 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 15 from Patent WO0194400.  
ACCESSION AX339033  
VERSION AX339033.1 GI:18129125  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 15 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source 1. .2021  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 547 a 406 c 462 g 605 t  
ORIGIN

```
Alignment Scores:
Pred. No.: 8,03e-60 Length: 2021
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-3 (1-131) x AX339033 (1-2021)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA 477

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 478 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAATATCAAGTATGGA 537

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 538 CGAGATGATGTGACTGATGAGAGATGGAGAGCAGCAAGGGAAGCAATCGGTATGAT 597

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 598 TTTATCATGAGGTTTCTTAATAAATTTAATACATTTGTTAGGGGAAAAAGAGCTCAATG 657

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsnProLysIle 100
DB 658 AGTGGAGGGCAGAAACAGAGGATCGCAATTCCTGTCGCTTAGTTCGAAACCCCAAGATT 717

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 718 CTGATTTTAGTAGAGGCTAGCTCTGCTCGCTTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

QY 121 AlaLeuGluLys 124
DB 778 GCACCTGGAGAAG 789

RESULT 5
AX339030 LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 12 from Patent WO0194400.
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1..3177
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2.15e-55 Length: 3177
Score: 619.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.93% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-3 (1-131) x AX339030 (1-3177)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 739 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA 798

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 799 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAATATCAAGTATGGA 858

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 859 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAATCGGTATGAT 918
```

```
Alignment Scores:
Pred. No.: 8,03e-60 Length: 2021
Score: 659.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-3 (1-131) x AX339033 (1-2021)

QY 1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
DB 1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTGGA 1323

QY 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
DB 1324 GTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAATATCAAGTATGGA 1383

QY 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
DB 1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAATCGGTATGAT 1443

QY 61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
DB 1444 TTTATCATGAGGTTTCTTAATAAATTTAATACATTTGTTAGGGGAAAAAGAGCTCAATG 1503

QY 81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsnProLysIle 100
DB 1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTCCTGTCGCTTAGTTCGAAACCCCAAGATT 1563

QY 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
DB 1564 CTGATTTTAGTAGAGGCTAGCTCTGCTCGCTTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623

QY 121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
DB 1624 GCACCTGGAAGAAGATACCCCGAGTATTCATTT 1656

RESULT 4
AX339028 LOCUS AX339028 2856 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 10 from Patent WO0194400.
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1..2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN

Alignment Scores:
Pred. No.: 1.9e-55 Length: 2856
Score: 619.00 Matches: 124
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.93% Indels: 0
DB: 6 Gaps: 0

US-09-873-409-3 (1-131) x AX339028 (1-2856)
```



RESULT 8  
AX339031  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT  
ORIGIN  
Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:  
US-09-873-409-3 (1-131) x AX339031 (1-3702)  
Qy 1 MetValAspGluHenAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGly 20  
Db 1264 ATGGTGGATGAGAAATGACATCAGACGCTTTAAATGTGGCGCATATATGAGACCATATTTGGA 1323  
Qy 21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGly 40  
Db 1324 GTGGTTAGTCAAGAGCTGTTTGTTCGGACCAACCATGATGATCAATATCAAGTATGGA 1383  
Qy 41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAsp 60  
Db 1384 CGAGATGATGCTGATGATGAAGAGATGAGAGAGCAGCAAGGGAAGCAATGCGTATGAT 1443  
Qy 61 PheIleMetGluPheProAenLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80  
Db 1444 TTTATCATGGAGTTTCCATAATTAATTAATGATGTTGGGGGAAAAGAGCTCAATG 1503  
Qy 81 SerGlyGlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAenProLysIle 100  
Db 1504 AGTGGAGGGGAGAAACAGAGATGCGAATTTGCTGCTGCTAGTTCGAAACCCCAAGATT 1563  
Qy 101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120  
Db 1564 CTGATTTAGATGAGGCTAGCTGCTGCTGCTGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623  
Qy 121 AlaLeuGluLys 124  
Db 1624 GCACCTGGAGAAG 1635  
RESULT 9  
AB029153  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT  
ORIGIN  
Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:  
US-09-873-409-3 (1-131) x AB029153 (1-3489)  
Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21  
Db 1180 ATTCATGGACAGACATTTAGGACCAATAATGTGAGATATCTTCGGGAAATCAATTGGTGTG 1239  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGlyArg 41  
Db 1240 GTGAGTCAGGAACCTGTGTGTTTGCACATACATAGCTGAGAAACATTCCTATGCGCGT 1299  
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe 61  
Db 1300 GAAATGTACCATGAGGAGGATGAGAAAGCTGTCAAGGAGCAATGCTATGACTTT 1359  
Qy 62 IleMetGluPheProAenLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
1  
Okai,Y., Nakamura,N., Matsushiro,H., Kato,H., Setoguchi,A.,  
Yazawa,M., Okuda,M., Watari,T., Hasegawa,A. and Tsujimoto,H.  
Molecular analysis of multidrug resistance in feline lymphoma cells  
Am. J. Vet. Res. 61 (9), 1122-1127 (2000)  
20431258  
2 (bases 1 to 3489)  
Tsujimoto,H. and Okai,Y.  
Direct Submission  
Submitted (05-JUN-1999) Hajime Tsujimoto, The University of Tokyo,  
Department of Veterinary Internal Medicine, The University of  
Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan  
(E-mail:atsujihongo.ecc.u-tokyo.ac.jp, Tel:+81-3-5841-8004,  
Fax:+81-3-5841-8178)  
FEATURES  
Location/Qualifiers  
1..3489  
/organism="Felis catus"  
/db\_xref="taxon:9685"  
/cell\_line="adriamycin-resistant subline of FT-1"  
/cell\_type="feline lymphoma cell line"  
<1..>3489  
/note="multi-drug resistance related mRNA"  
/codon\_start=1  
/protein\_id="BAA87071.1"  
/translation="HGAALPLMLVFGDMTDFANAGISRNLTTLNITGESIVNDSYF  
INRLEEMTTVAYYSGIGAGVLAAYIQVSPFCLAAGRIILKIRKQFPFHAMRQEVG  
WPDVHDVGLNTRLTDDVSKINEGIDKIMFQSMATFIFIGVFTGKWLTLVIL  
ALSPVLGLSAAIWAIIKASPTDKELLAYAKAGAAVEVLAARTVIAFGQKELERY  
NKNLEAKRIGIKKAITANISIGIAFLIYASIALAFWYGTSLVLSHEYSIQGLTVF  
FSVLGAFSVGQASPIEAFANARGAAVEIFKIDNKPISIDSYKNGHKPDNKGLE  
FKNVHFSYPSRKEVILKGLNLKVGSGGTVALVGNSSCGKSTTVQLMORLYDPTGMV  
SIDGQDIRTNVYRLREIIGVVSQBPVLFAITAEINRYGRENVTMEETKAVKEANA  
YVIVKLPNKPTLVGERGAOLSGGOKORIAARALVRNPKILLDEATSDSESEA  
VVOVALDKARKGRTIIVVAHRLSTIRNADVIAGFDGVIKEGNHDELMKESGIVFKL  
VTMQTRNEIELEENAVYESISEIDALEMSKDSGSLIRRRSTRKSIHASQODRKLG  
TKENLDERVPVPSFWRLKLNITENPFFVFGICAIINGQLQPSFVILSIGVFTFR  
VEDPTEKQNSIFSLFLGLGISITPFLQGTFGKAGEILTTLKRLTVFRSMRLQ  
DYSWDPDKNTGALTTLANDAAQVGAIGSLRAITQNIANLGTGIIISLIYGNQL  
TLLALIVPIIAIAGVEMKLSGOALKDKKEGAKGVATEIENFRVTVSLTREQK  
FEYMTAQSLQVPYRNSLRKAHIFGISFSTIQAMYPYSYAGCFRFGALVAHEPMDQD  
VLVLSFSAIVFGAMVQVSSFPDYAKAVSAHVIMILEKIPLDISTYSTGEMPNIL  
EGNVTFEYMFNPTRPDIPVLQGLSLEVKGTTLALVGGSGGKSTVOLLERFYDP  
MAGTVLIDCKEIKHLNVOLWRMGIVSQEPLFDICIGENIAYGNSRVVSQEEIVR  
AAKEANIHPFIETLPKYNTRVGDGTQSGGOKORIAARALVRQPRILLDEATSA  
LDTSEKVVQOEALDKAR"  
BASE COUNT 1015 a 716 c 847 g 911 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3..95e-40 Length: 3489  
Score: 474.00 Matches: 92  
Percent Similarity: 90.24% Conservative: 19  
Best Local Similarity: 74.80% Mismatches: 12  
Query Match: 71.93% Indels: 0  
DB: 4 Gaps: 0  
US-09-873-409-3 (1-131) x AB029153 (1-3489)  
Qy 2 ValAspGluAsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyVal 21  
Db 1180 ATTCATGGACAGACATTTAGGACCAATAATGTGAGATATCTTCGGGAAATCAATTGGTGTG 1239  
Qy 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGlyArg 41  
Db 1240 GTGAGTCAGGAACCTGTGTGTTTGCACATACATAGCTGAGAAACATTCCTATGCGCGT 1299  
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAenAlaTyrAspPhe 61  
Db 1300 GAAATGTACCATGAGGAGGATGAGAAAGCTGTCAAGGAGCAATGCTATGACTTT 1359  
Qy 62 IleMetGluPheProAenLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81



```

Db 1436 GTGAGTCAGAACCCGTCGTTGGCCACACCATTCGCCAAACATTCGCTATGCCGA 1495
QY 42 AspAspValThrAspGluMetGluArgAlaAalaArgGluAlaAalaValThrAspPhe 61
Db 1496 GAAACGTCACCATGATGAGATAGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTC 1555
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1556 ATCATGAATGCGCCACAAATTAACACCTGTTGGTGAGAGAGGGCGCAGCTGAGT 1615
QY 82 GlyGlyGlnLysGlnArgGluAlaAalaArgAlaValThrAspPheGlyAlaGlnMetSer 101
Db 1616 GGGGACAGAAACAGAGGATCGCCATTCGCCGGCCCTGTGTCGCAACCCCAAGATCCCT 1675
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAala 121
Db 1676 TTGTTGATGAGCGCCACGTCAGCTTGGACACAGAAAGCGGTCGTTGAGCGCGCT 1735
QY 122 LeuGluLys 124
Db 1736 CTGATAAG 1744

RESULT 12
AF099732
LOCUS
DEFINITION
Fundulus heteroclitus multidrug resistance transporter homolog
(mdr) mRNA, partial cds.
ACCESSION
AF099732
VERSION
AF099732.1 GI:4574223
KEYWORDS
Fundulus heteroclitus.
ORGANISM
Fundulus heteroclitus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
Cyprinodontiformes; Fundulidae; Fundulus.
REFERENCE
1 (bases 1 to 2752)
Cooper,P.S., Van Veld,P.A. and Reece,K.S.
P-glycoprotein related sequences from the mummichog (Fundulus
heteroclitus)
Unpublished
REFERENCE
2 (bases 1 to 2752)
Cooper,P.S., Van Veld,P.A. and Reece,K.S.
Direct Submission
Submitted (19-OCT-1998) School of Marine Science, Department of
Environmental Sciences, Virginia Institute of Marine Science,
College of William and Mary, Gloucester Point, VA 23062, USA
Location/Qualifiers
1. .2752
/organism="Fundulus heteroclitus"
/db_xref="taxon:8078"
/tissue_type="liver; intestine"
<1. .2752
/genes="mdr"
/genes="mdr"
/codon_start=1
/product="multidrug resistance transporter homolog"
/protein_id="AAD23956.1"
/db_xref="GI:4574224"

FEATURES
source
gene
CDS

```

```

OIGIVSQPVLFDCALENIAYGDSRVTMEETEAARAKAANIHFINELPQKYNTQA
GDVQLSGGQRIATAIARLNRNPKVLLDETSALDTESEKVVQDLDQSKGRTIC
IVGARLSTIRNADRIAVFQGGVVVEQTHOOLAKGVYHMLVTLQHGHE"

BASE COUNT 661 a 718 c 791 g 582 t
ORIGIN

Alignment Scores:
Pred. No.: 1e-39 Length: 2752
Score: 469.00 Matches: 91
Percent Similarity: 89.43% Conservative: 19
Best Local Similarity: 73.98% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 5 Gaps: 0

US-09-873-409-3 (1-131) x AF099732 (1-2752)
QY 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyVal 21
Db 58 ATCGATGGGACGACATCGCTCTTATATGCGGTACTGAGGGGAATGATCGCGTG 117
QY 22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLysTyGlyArg 41
Db 118 GTGAGCAGGAGCCCATCTCTTCCGCCACCATCGCTGAGAACATCGCTACGCGCA 177
QY 42 AspAspValThrAspGluMetGluArgAlaAalaArgGluAlaAalaValThrAspPhe 61
Db 178 CCGACGTCGACGAGGAGAGATCGAAAGCTGCCAAGGAGCCCAATGCTTACGACTTC 237
QY 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 238 ATCATGAACCTTCTGTATAGTTTGAGACACTAGTCGGGACCGGAGAACCCAGATGAGC 297
QY 82 GlyGlyGlnLysGlnArgIleAlaIleArgAlaLeuValArgAsnProLysIleLeu 101
Db 298 GAGGACAGACAGCAGAGGATTCGATCGCTCGAGCTTGGTCCGAAACCCAAATCCTG 357
QY 102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAala 121
Db 358 CTGTTGAGCAAGCCACGCTGCTCCCTGGATGCTGAGAGTGAGACTATCGTTCAAGCTGCA 417
QY 122 LeuGluLys 124
Db 418 CTCGATAAG 426

RESULT 13
CFA419568
LOCUS
DEFINITION
Canis familiaris mRNA for abortive P-glycoprotein (p-gp gene).
ACCESSION
AJ419568
VERSION
AJ419568.1 GI:17385398
KEYWORDS
P-glycoprotein; p-gp gene.
SOURCE
dog.
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1
Roulet,A.
Characterization of a MDR1a P-gp deficient dog in relation to
ivermectin sensitivity
Thesis (2002) Department of Biological Sciences, Universite de
Toulouse (UPS III), Toulouse, France
REFERENCE
2 (bases 1 to 3934)
Roulet,A.
Direct Submission
Submitted (07-NOV-2001) Roulet A., Pharmacology, Inra, 180 Chemin
de Tournnefeuille, 31931 Toulouse cedex, FRANCE
Location/Qualifiers
1. .3934
/organism="Canis familiaris"
/db_xref="taxon:9615"
/tissue_type="liver"
/notes="Ivermectin sensitive collies"
7. .282
gene

```



```

CDS
/genes="p-gp"
7..282
/genes="p-gp"
/note="abortive product"
/codon_start=1
/evidence=experimental
/product="P-glycoprotein"
/protein_id="CAD12806.1"
/db_xref="GI:17385399"
/translation="MDPEGRGSAEFKNFKMKKSKNEKKKPKTPTVSTFAMPRYSN
WDLRLVNLVGTWAAIHGAALPLMLVFGNMTALQMDFQETKLFQL"
BASE COUNT 1144 a 815 c 964 g 1011 t
ORIGIN

Alignment Scores:
Pred. No.: 1 52e-39 Length: 3934
Score: 469.00 Matches: 90
Percent Similarity: 89.43% Conservative: 20
Best Local Similarity: 73.17% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 4 Gaps: 0

US-09-873-409-3 (1-131) x CFA419568 (1-3934)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1365 ATTGATGCAGAGACATTAGGACCATTAATGTAAGGCATCTTCGGGAAATTAAGTGGTGTG 1424
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1425 GTGAGTCAGAGCGCTGTGTGTTTGCACACAGCATAGCTGAAAAACATTCGCTATGGCCGC 1484
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1485 GAAATGTCACCATGATGAGATGAGAGAGCTGTAAGAGCAATGCTATGATTTT 1544
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db 1545 ATCATGAACTACCTAATAATTTGACACTCTGTTGGAGAGAGAGGGCCAGCTGAGT 1604
Qy 82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db 1605 GGTGGACAGAAACAGAGAAATCGCCATTCCTGGGCCCTGTTGTCACACCCCAAGATTCCT 1664
Qy 102 IleLeuAspGluAlaThrSerAlaLeuAspSerAlaSerLysSerAlaValGlnAlaAla 121
Db 1665 CTGCTGATGAGCAACATGCACTCTGACACTCTGGAAGTGAAGCAGTGTTCAGGTGCC 1724
Qy 122 LeuGluLys 124
Db 1725 CTGGATAAG 1733

RESULT 14
GGA9799
LOCUS Gallus gallus mRNA for ABC transporter protein. VRT 03-MAR-1999
DEFINITION Gallus gallus mRNA for ABC transporter protein.
ACCESSION AJ009799
VERSION AJ009799.1 GI:3355756
KEYWORDS ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 4018)
Edelmann,H.M.L.
Direct Submission
Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular
Genetics, Bioenter and University of Vienna, Dr. Bohr-Gasse 9/2
A-1030 Vienna, A-1030 Vienna, AUSTRIA
2 (bases 1 to 4018)
Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C.,
Kane,S.E. and Kuchler,K.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
```

```

TITLE Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
interacts with estradiol
JOURNAL Biol. Chem. 380 (2), 231-241 (1999)
MEDLINE 99209805
PUBMED 10195430
FEATURES
source
Location/Qualifiers
1..4018
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="intestine"
109..3975
/genes="cmdr1"
109..3975
/genes="cmdr1"
/note="ABC transporter protein"
/codon_start=1
/product="P-glycoprotein"
/protein_id="CAA08835.1"
/db_xref="GI:3355757"
/db_xref="SPTREMBL:O93437"
/translation="MSEDEKQKHTVDGNTYIEIANSNQDQDPEDKKKKKKKKPKQM
VSLALFRYSCTDKLIMFGSLAIAHGTSLPIAMIIFGDMTDSFVTSQMTNITGNS
SGLNSSADVENKLEEMTRYAYVYSAIAAALVAAVYIOTSPFTLAAGQVKIRKFF
HAIHQEIGCFDNDAGELNTRLDDVSKINEGIGDKTGLTQSETTLTGIVGFIR
GKMLTLVAVSPVLGSLAALWAKLLTAFTDKQAYAKAGAVAEVLSAVRTVAPG
QOEIKRYHNLEDAKIGIRKAITSNISMGAAFLLIYASVALAFWYGTTLILANEY
STGNLTVFVSLIGAFSIGOTAPSIEAFANAGAAAYAFNIIDNEPIDSADGKH
PDHKNLEFQNVFNPSPDVEILKGLNKVNCQGTVALVGGCCGKSTTVQIQR
FYDPKEGTTIDGDLKSLNRYLREIIGVNOBPVLFTTIAENIRYGREVDTMEEI
ERATKANAYDFIMKLPKFTVCGERGAQSGGQKQRIATARALVHNPKILLDEAT
SALDTESBVQAALDKAREGRTVVVAHRLSTVNRADLIAVFESGVITEQGNHSQLI
EKGIYKLVNMQTIEDPSSEKSNVSKVSGSNLDESLLKELRRGSTRSMK
KPEPNDDEKSGSPDELPVFLKMLKNEMPFYVAGTFCAIVKAGALQAFSVI
PSEIIGIFSETDOKVRLKSLNLSLLFLALGIIISFTFVQGFAGKAGIITMKLRF
WAFKMLRQDMAPDDPKNSTCALTRLANDASOVKGATGRLALIAONIANLGTGII
LSLVYRWQTLTLLAVPIIAVAGMIEMLAGHAKDKLELEPAKATATEAENIRT
VASLTKRERFELMYGERLLVYRNSVKAHIFGFCLSQAMFTYACCFRPGALV
VNGHTYEYTVLFAVSADVAGALQTSFAPDYAKAKISAHLFVLFNRVPIIDSIR
EDGEKPEKFGNTRIKDYKNYPNRPVKILQGLNLAVEKGETLALVFNSSGCGTIV
OLLERFYDPLSGEIVFDIDIAKTLNIQWLRIHIGIVSOEPIFDFTIARANIAYGNSR
EVSHEIISAAKAASIHFDISLPEKYNTRVGDGKTGSLSGGOKQRIATARALIRKPOI
LLLDATSAIDTESKIVQELADKAREGTCIVIAHRLSTIQNADKIAVIQNGKVIQ
GTHQQLLAEKGFYISLVNVQSGCNN"
mat_peptide
112..3972
/genes="cmdr1"
/product="ABC transporter protein"
/note="P-glycoprotein"
BASE COUNT 1288 a 758 c 911 g 1061 t
ORIGIN

Alignment Scores:
Pred. No.: 1 56e-39 Length: 4018
Score: 469.00 Matches: 91
Percent Similarity: 89.43% Conservative: 19
Best Local Similarity: 73.98% Mismatches: 13
Query Match: 71.17% Indels: 0
DB: 5 Gaps: 0

US-09-873-409-3 (1-131) x GGA9799 (1-4018)
Qy 2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db 1489 ATTGACGGGAGGATCTGAAGAGCCTGAATGTAAGATATATCGCGAGAGATTATGGTGTG 1548
Qy 22 ValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArg 41
Db 1549 GTGAACAGAGAGCCGCGTGTGTTGCTACAACTATTGCAAAATATTCGTTATGGCCGT 1608
Qy 42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db 1609 GAGGATGTCACCATGGAAGAATTTGAAGAGCTACCAAGGAGCTAATGCTTATGATTC 1668
Qy 62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 7277.11 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-4  
Perfect score: 5330  
Sequence: 1 MVISITKSLAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115420\_19240/app\_query.fasta\_1.7544  
-DB=EST -QMR=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1.1 30544 @runat\_27032003\_115420\_19240 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	20.5	669	13	BM013981	BM013981 603639619
2	1051	19.7	1033	14	BM926413	BM926413 AGENCOURT
3	1008	18.9	943	9	AL520322	AL520322 AL520322
4	995	18.7	760	13	BM016204	BM016204 603642659
5	918	17.2	545	9	AL040762	AL040762 DRFP434C
6	903	16.9	929	14	BQ882401	BQ882401 AGENCOURT
7	856.5	16.1	948	13	BM471690	BM471690 AGENCOURT
8	834	15.6	998	14	BM904842	BM904842 AGENCOURT
9	780	14.6	2676	11	AK014319	AK014319 Mus muscu
10	771	14.5	894	12	BF584668	BF584668 602098406
11	770.5	14.5	1019	12	BG248052	BG248052 602359987
12	753	14.1	726	12	BG293345	BG293345 602390738
13	752.5	14.1	871	17	AZ682350	AZ682350 ENTIL96TF
14	751	14.1	944	12	BF796582	BF796582 602258463
15	743.5	13.9	886	17	BH139685	BH139685 ENTNA47TR
16	734.5	13.8	932	17	AZ670821	AZ670821 ENTUN69TF
17	728.5	13.7	947	17	AZ683753	AZ683753 ENTIL96TF
18	726.5	13.6	886	17	AZ540627	AZ540627 ENTQ18TF
19	723.5	13.6	880	17	AZ687805	AZ687805 ENTIL96TF
20	721.5	13.5	939	14	BQ720763	BQ720763 AGENCOURT
21	720.5	13.5	897	17	AZ541090	AZ541090 ENTDS67TR
22	719	13.5	780	12	BG587938	BG587938 EST489713
23	711.5	13.3	913	17	BH155700	BH155700 ENTRO54TR
24	709.5	13.3	834	17	AZ548312	AZ548312 ENTIF67TR
25	708.5	13.3	853	17	AZ679807	AZ679807 ENTH167TR
26	705.5	13.2	823	17	AZ532602	AZ532602 AGENCOURT
27	697.5	13.1	861	14	BQ717101	BQ717101 AGENCOURT
28	692.5	13.0	1341	11	AI108285	AI108285 Zea mays
29	690.5	13.0	899	17	BH154857	BH154857 ENTRO47TF
30	690	12.9	835	13	BJ349604	BJ349604 BJ349604
31	689.5	12.9	891	17	AZ682250	AZ682250 ENTKT68TR
32	688	12.9	1123	11	AV108485	AV108485 Zea mays
33	687	12.9	715	14	BQ869512	BQ869512 QGD6G09.Y
34	684	12.8	820	13	BJ356540	BJ356540 BJ356540
35	683	12.8	600	13	BM486593	BM486593 Pgm2n.pk0
36	682.5	12.8	919	17	AZ690701	AZ690701 ENTH168TR
37	680	12.8	926	14	BQ123477	BQ123477 EST609053
38	677.5	12.7	795	12	BF113560	BF113560 601900192
39	674.5	12.7	815	13	BI409057	BI409057 602961127
40	670	12.6	852	17	BH720383	BH720383 BOH2294TF
41	669	12.6	547	13	BJ030711	BJ030711 BJ030711
42	667	12.5	937	17	BH150760	BH150760 ENTQ42TR
43	666	12.5	792	13	BJ352404	BJ352404 BJ352404
44	663	12.4	838	13	BJ355307	BJ355307 BJ355307
45	662	12.4	636	10	AV962688	AV962688 AV962688

ALIGNMENTS

RESULT 1  
BM013981  
LOCUS 603639619F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5415556 5', linear EST 30-OCT-2001  
DEFINITION mRNA sequence.  
ACCESSION BM013981  
VERSION BM013981.1 GI:16528335  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 669)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



QY 169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188  
 |||||  
 Db 300 AGTATAGATAACCTTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTGGAA 359  
 |||||  
 QY 189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208  
 |||||  
 Db 360 TTTTAAATATGTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGTCTG 419  
 |||||  
 QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228  
 |||||  
 Db 420 AATCTCAGAAATTAAGTCTGGAGACAGCTCGCTTGTGCTCTCAATGGCATGGGAAG 479  
 |||||  
 QY 229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248  
 |||||  
 Db 480 AGTACGGTAGTCCAGCTTCTGCAGAGCTTATATGATCCCGATGATGGCTTATCATGGTG 539  
 |||||  
 QY 249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268  
 |||||  
 Db 540 GATGAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTTGAGTGGTT 599  
 |||||  
 QY 269 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 288  
 |||||  
 Db 600 AGTCAAGAGCTGTTTGTTCGGACACCATCAGTACAAATATCAAGTATGACGAGAT 659  
 |||||  
 QY 289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 308  
 |||||  
 Db 660 GATGTGACTGATCAAGAGATGGAGAGACGACGAAGGGAACCAATGCGTATGATTTATC 719  
 |||||  
 QY 309 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 328  
 |||||  
 Db 720 ATGGAGTTTCTCTAATAA-TTTAATACATTGGTAGGGGAANAAGAGCTCCAATCAGTGA 778  
 |||||  
 QY 329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348  
 |||||  
 Db 779 AGCCAGAACCG-AGGATCCCAATTGCTCTCGCTTAGTTCGAAACCCAGGAT-CTGAAT 836  
 |||||  
 QY 349 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 364  
 |||||  
 Db 837 TTAGATGAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884  
 |||||  
 RESULT 3  
 AL520322 943 bp mRNA linear EST 13-FEB-2001  
 LOCUS AL520322 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB06YC15 5  
 DEFINITION Prime, mRNA sequence.  
 ACCESSION AL520322  
 VERSION AL520322.1 GI:12783815  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 943)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 FEATURES  
 source  
 1..943  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB06YC15"  
 /clone\_lib="LTI\_NFL004\_NBC2"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed  
 by Life Technologies. Contact : Feng Liang Life  
 Technologies, a division of Invitrogen 9800 Medical Center  
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610  
 8371 Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t 1 others  
 ORIGIN

Alignment Scores: 7.11e-105 Length: 943  
 Pred. No.: 1008.00 Matches: 134  
 Score: 1008.00  
 Percent Similarity: 80.40% Conservative: 48  
 Best Local Similarity: 64.45% Mismatches: 58  
 Query Match: 18.91% Indels: 1  
 DB: 9 Gaps: 0

US-09-873-409-4 (1-1058) x AL520322 (1-943)

QY 757 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 776  
 |||||  
 Db 5 TTATTATTTTTCAGCTGTTGCTTTGTGGCCATGGCGTGGGCAAGTCAGTTTCATTGCT 64  
 |||||  
 QY 777 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 796  
 |||||  
 Db 65 CTGACTATGCCAAAGCCAAAATATCAGCAGGCCACATCATCATGATCATTTGAAAAAAC 124  
 |||||  
 QY 797 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 816  
 |||||  
 Db 125 CCTTTGATTGACAGCTACAGCAGCGAAGGCGCTAATGCCGAACACATTTGGAAGAAATGTC 184  
 |||||  
 QY 817 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 836  
 |||||  
 Db 185 ACATTTGGTGAAGTTGTATTCATATCCACCAGCCGACATCCAGTGTCTCAGGA 244  
 |||||  
 QY 837 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 856  
 |||||  
 Db 245 CTCAGCCTGGAGGTGAAGAAGGCCAGACGCTGGCTCTGGTGGGCGAGCAGTGTGCTGGG 304  
 |||||  
 QY 857 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 876  
 |||||  
 Db 305 AAGAGCACATGTGCTCAGCTCTCGAGCGGTTCTACAGCCCTTGGCAGGGAAGTGTCTG 364  
 |||||  
 QY 877 PheAspGlyValAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 896  
 |||||  
 Db 365 CTGTATGGCAAAAGAAATTAAGCCACTGAATGTTCACTGGCTCCGAGCACACCTGGGCATC 424  
 |||||  
 QY 897 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 916  
 |||||  
 Db 425 GTGTCCCGAGGAGCCATCCTGTTGACTGTCAGCATTTGCTGAGAACATTTGCTATGGAGAC 484  
 |||||  
 QY 917 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 936  
 |||||  
 Db 485 AACAGCCGGGTGGTGTCCACAGGAAGAGATCGTGAGGCGACAAAGAGGCAACATACAT 544  
 |||||  
 QY 937 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 956  
 |||||  
 Db 545 GCCTTATCAGTCACTGCTCTTAATTAATATAGCATTAAAGTAGGAGACAAGGAAGTCTAG 604  
 |||||  
 QY 957 LeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 976  
 |||||  
 Db 605 CTCTCTGGTGGCCAGAAACAACGCAATTTGCCATAGCTGCTGCTGCTGCTGCTGCTGCT 664  
 |||||  
 QY 977 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 996  
 |||||  
 Db 665 ATTTTGTCTTTTGGATGAAGCCACGCTCTGATACACAAAAGTGAAGGTTGTCTCAA 724  
 |||||  
 QY 997 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1016  
 |||||  
 Db 725 GAAGCCCTGCAGAAAGCCAGAGAAGGCCGACCTGCATGTGATGTGTGCTACCGCTGTCC 784  
 |||||  
 QY 1017 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGln 1036

```

|||||
Db 785 ACCATCCAGATGACAGCTTAATAGTGTGTCTTTTCAGAAATGCGCAGAGTCAAGGACATGG 844
|||||
Qy 1036 yThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 1056
|||||
Db 845 CACGCATCAGCAGCTGCTGGCACAGAAAGGATCTATATTTTCAATGGTCAAGTGTCCAGGC 904
|||||
Qy 1056 r 1056
Db 905 T 905

RESULT 4
BM016204 760 bp mRNA linear EST 30-OCT-2001
LOCUS 603642659P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
DEFINITION mRNA sequence.
ACCESSION BM016204
VERSION BM016204.1 GI:16530558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-xemail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM12068 row: e column: 16
High quality sequence stop: 740.
FEATURES
Location/Qualifiers
source 1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1..55e-103 Length: 760
Score: 995.00 Matches: 203
Percent Similarity: 98.07% Conservatives: 0
Best Local Similarity: 98.07% Mismatches: 2
Query Match: 18.67% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-4 (1-1058) x BM016204 (1-760)

Qy 166 LysLysProSerIleAspAsnPhSerThrAlaGlyTyrLysProGluSerIleGluGly 185
Db 3 AAGAAACCCAGTATAGTAACCTTTTCCACAGCTGGATATAAACCCTGAATCCATGAAGGA 62
Qy 186 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 205
Db 63 ACTGTGGAATTTAAATAGTTCTTTCAATATATCATCAAGACCATCTATCAAGATCTG 122
Qy 206 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 225

```

```

|||||
Db 123 AAAGGTCGTAATCTCAAGATTAAGTCTGGAGAGACAGTCGCCCTTGGTCGCTCAATGGC 182
|||||
Qy 226 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 245
|||||
Db 183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTCGACAGGTTATATATCCGGAATGATGGCTTT 242
|||||
Qy 246 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 265
|||||
Db 243 ATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302
|||||
Qy 266 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 285
|||||
Db 303 GGAGTGGTTAGTCAAGAGCCCTGTTTGTTCGGACCAACCATCATAGTAACAATATCAAGTAT 362
|||||
Qy 286 GlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 305
|||||
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGGAAGCAATGCGTAT 422
|||||
Qy 306 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 325
|||||
Db 423 GATTTATCATGAGTTCTCTAATAATTTAATACATTTGGTAGGGGAAAAGAGCTCAA 482
|||||
Qy 326 MetSerGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 345
|||||
Db 483 ATGAGTGGAGGGCAGAAACAGAGATCGCAATTTGGCTCGTCTAGTTGGAACCCCAA 542
|||||
Qy 345 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 365
|||||
Db 543 GATTCGTATTTTCAGATGAGGCTACGTCCTCCCTGGATTGAGAAAGCAAGTCAGCTGTC 602
|||||
Qy 365 InAlaAlaLeuGluLys 370
|||||
Db 603 AAGCTGCACCTGGAGAG 619

RESULT 5
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKF2p434C1815 r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKF2p434C1815 5', mRNA sequence.
ACCESSION AL040762
VERSION AL040762.1 GI:5409708
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Bloecker H
MIPS
Am Klopferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKF2p434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source 1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"

```



```
/lab_host="DH10B"
BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN
Alignment Scores:
Pred. No.: 6,58e-95 Length: 545
Score: 918.00 Matches: 181
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.22% Indels: 0
DB: 9 Gaps: 0
US-09-873-409-4 (1-1058) x AL040762 (1-545)
QY 169 SerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
Db 3 AGTATAGATTAACATTTCCACAGCTGGATATACCTGAATCCATAGAGCACTGTGGAA 62
QY 189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
Db 63 TTTAAATAATGTTCTTTCAATATCCATCAAGACCATCTCAAGATTCTGAAAGGTCTG 122
QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
Db 123 AATCTCAGAATTAAAGTCTGGAGACAGTCGCTTGGTCTGCTCAATGGCAGTGGGAAG 182
QY 229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
Db 183 AGTACGGTAGCTCAGCTTCTGCAGAGGTTATATGATCCGATGATGCTTTATCATGTTG 242
QY 249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268
Db 243 GATCAGAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATTTGAGTGGT 302
QY 269 SerGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 288
Db 303 AGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACATATCAAGTATGACGAGAT 362
QY 289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnLatyrAspPheIle 308
Db 363 CATGTGACTGATGAAGATGGAGAGCAGCAGGAGCAAGGAGCAATGCGTATGATTTATC 422
QY 309 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 328
Db 423 ATGGAGTTTCTTAATTAATTAATATACATTGGTAGGGGAAAAAGAGCTCAATGAGTGA 482
QY 329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348
Db 483 GGGCAGAAACAGAGGATCGCAATTGCTCGCTTAGTTCGAAACCCCAAGATTCTGATT 542
QY 349 Leu 349
Db 543 TTA 545
RESULT 6
LOCUS BQ882401
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
5', mRNA sequence.
ACCESSION BQ882401
VERSION BQ882401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mai.nih.gov
```

```
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM2492 row: c column: 15
High quality sequence stop: 677.
Location/Qualifiers
1. .929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:6291782"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"
BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 8,72e-93 Length: 929
Score: 903.00 Matches: 188
Percent Similarity: 73.02% Conservative: 61
Best Local Similarity: 55.13% Mismatches: 58
Query Match: 16.94% Indels: 34
DB: 14 Gaps: 3
US-09-873-409-4 (1-1058) x BQ882401 (1-929)
QY 12 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThr-ValI1 31
Db 1 GCATATGCAAAAGCTGGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAACTGGTAT 60
QY 31 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 51
Db 61 TGCATTGGAGGACAAAGAAAGAACTTGAA----- 91
QY 51 rAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrTh 71
Db 92 -----AGGTACAA 99
QY 71 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 91
Db 100 CAAAATTTAGAGAGAGCTAAAGAAATGGGATTAAGAAAGCTATTACAGCCAAATTTC 159
QY 91 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 111
Db 160 TATAGGTGCTGCTTCTCTGCTGATCTATGCTATCTATGCTCTGCTCTGCTATGGAC 219
QY 111 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 131
Db 220 CACCTTGGTCTCTCAGGGGAA-----TATCTATTTGGACAAAGTACTACTGTATTCTT 273
QY 131 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 151
Db 274 TTTCTGATTAATTTGGGGCTTTTAGTGTGGACAGGAGCTCTCCAGGCAATTGAAGCAATTGC 333
QY 151 aIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 171
Db 334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTAATAAGCAAGTATTGA 393
QY 171 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 191
Db 394 CAGCTATTGCAAGAGTGGGCACAAACCAAGATATATTAAGGGAAATTTGGAATTCAGAA 453
```

```
FEATURES
source
1. .929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:6291782"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"
BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN
```

QY 191 nValSerPheAsnTyrProSerArgProSerIleLeuLysGlyLeuAsnLeuAr 211  
 Db 454 TGTCACCTTCAGTACCACTCGAAGAAGTTAAGATCTTGAAGGCTCTGAACCTGAA 513  
 QY 211 gileLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 231  
 Db 514 GGTGCAGAGTGGGAGACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAA 573  
 QY 231 lValGlnLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 251  
 Db 574 AGTCACGCTCATGCAGAGGCTCTATGACCCACAGAGGGGATGTCAGTGTGTATGGACA 633  
 QY 251 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 271  
 Db 634 GGATATTAGACCAATAATGTAAGTTTCTACGGGAATCATTTGGTGTGTGAGTCAGGA 693  
 QY 271 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValTh 291  
 Db 694 ACCTGTATTGTTCCACACAGTAGTGAACATTTCCCTATGGCCGTGAATAATGTCCAC 753  
 QY 291 rAspGluGluMetGluArgAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 311  
 Db 754 CATGGATGAGATTGAGAAAGCTGTCAGGAAGCAATGCCCTATGACTTTATCATGAAC 813  
 QY 311 heProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGly---G 330  
 Db 814 TGCCTCATTAATTTGACACCTGTTGGANNAAGAGGGCCCAATTCAGTGGTGGGCGAC 873  
 QY 330 lnLysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 347  
 Db 874 AACCAAAAGGATCTCCCTGGGACGTGGCCCTGTTCCCAACCCCAAGATCCTC 928

RESULT 7  
 BM471690 948 bp mRNA linear EST 05-FEB-2002  
 LOCUS AGENCOURT\_6465349 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5539117  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM471690  
 VERSION BM471690.1 GI:18520732  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rgs@bbs-e@mail.nih.gov  
 Tissue Procurement: ATCC/DCTP/DTP  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM12233 row: b column: 14  
 High quality sequence stop: 569.  
 Location/Qualifiers  
 1..948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5539117"  
 /clone\_lib="NIH\_MGC\_72"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 Kb. Library constructed by Life Technologies."  
 FEATURES source  
 285 a 187 c 194 g 282 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,06e-87 Length: 948  
 Score: 856.50 Matches: 186  
 Percent Similarity: 89.67% Conservative: 5  
 Best Local Similarity: 87.32% Mismatches: 14  
 Query Match: 16.07% Indels: 8  
 DB: 13 Gaps: 1

US-09-873-409-4 (1-1058) x BM471690 (1-948)

QY 129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 148  
 Db 205 GTTTCTTTAGTGAATCCATAGCAGTATTCATTTGGAGCAGCAGTCCCTCACTTGA 264  
 QY 149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 168  
 Db 265 ACCTTCGCAATAGCCGAGGAGCTGCTTTTCATATTTCCAGGTATTGTATAAGAACCC 324  
 QY 169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188  
 Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAGGAACCTGTGAA 384  
 QY 189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208  
 Db 385 TTTAAAAATGTTCTTTCAATATTCATCAAGACCATCTATCAAGATTTCTGAAAGGTCTG 444  
 QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228  
 Db 445 AATCTCAGAAATTAAAGTCTGGAGAGCAGTCGCTTGGTGGTCTCAATGCGAGTGGGAAG 504  
 QY 229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248  
 Db 505 AGTACGGTAGTCAGCTTCTGCAGAGTTATATGATCCGATGATGCTTTATCATGTTG 564  
 QY 249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 268  
 Db 565 GATGAGAATGACATCAGAGCTTTTAAATGTGCGCATTTATCGAGACCATATTTGGAGTGT 624  
 QY 268 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 287  
 Db 625 TAGTCAAGAGCCCTGGTTTGTTCGGACCCACCATCAGTAAACAATATCAAGTATGGGACG 684  
 QY 287 GAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 306  
 Db 685 AGATGATGGAGCTGTGAACAAATGGAAGAACCAACAGGGAACCAAGGCGAATGAA 744  
 QY 307 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 325  
 Db 745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTTGGAAGGGGAAAAAGGACCTTCA 804  
 QY 325 nMetSerGlyGlyGlnLysGlnArg 333  
 Db 805 AGGAATGGAAGGGCCCAAAACCCCAAG 829

RESULT 8  
 BM904842 998 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOURT\_6699581 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5557655  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM904842  
 VERSION BM904842.1 GI:19355221  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 998)  
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.

NC, CAP Trapper.	Mus musculus (Strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library clone:3222401P09.
Mus musculus	
1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	
AUTHORS	Carninci,P. and Hayashizaki,Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	93279253
PUBMED	10349636
2	
REFERENCE	
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh M., Konno H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
3	
REFERENCE	
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kagiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
4	
REFERENCE	
AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotjohori,T., Bono,H., Kasukawa,T., Saito,R.,



```

792 GTGCAGAAATTCAGTATCATCTTCAAGCTTGTCTCATCTCCAGGAGCTGCGCAGCTGCACC 851
623 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 642
852 CAGGTGATGGTAGCTGGTGTCCCTGTCTATGCTGTCCCTCCCTTACCTTACCTGATGCTG 911
643 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 662
912 GCTGTGTCACACCCGCTCATGGGAGTGGGCACCTGTATGGCTCAGGCTCCGAAAG 971
663 PheAlaAsnIleAspIleGlnLeuLeuLeuHisAlaGlyIleIleAlaThrGluAlaLeu 682
972 CTCTCTCGCAGCAGTGTGAGGAGCATGTCAGGCGCAACAGGTGTAGCAGATGAGGCCCTT 1031
683 GluAsnIleArgThrIleValSerLeuThrArgGluIleAlaPheGluGlnMetTyrGlu 702
1032 GGCATGTTCCGAGCTGTGCGGGCTTCGCCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1091
703 GluMetLeuGlnThrGlnHisArgAsnThrSerIleIleGlnIleIleGlySerCys 722
1092 GCAGAACTGGAG-----TCATGCTGCTGTAAGCAGAGCAAGCACTGGCGCAGGGC 1139
723 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyr-----AlaAlaGlyPheArg 739
1140 ATCGCTTGTTCAGAGGCTCTCCACATCGCTTCAACTGTATGCTTCTGGGCAACCTG 1199
740 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly-----MetPhe 756
1200 TTCATTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAAGGGGAGACCTCATGTCTCTC 1259
757 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyIleThrLeuValLeuAla 776
1260 CTGGTGGCTTCCAGACAGTACAGAGCTATGGCC-----AGCCTCTGTGCTCTGTTT 1313
777 ProGluTyrSerIleAlaIleSerGlyAlaAlaHisLeuPheAlaLeuLeuGluIleLys 796
1314 GGTCAAGTGTGAGTGGCTGAGTGGCGGAGCCGAGCTCTTCGAATACATGCGCCCTGAGC 1373
797 ProAsnIleAspSerArgSerGlnGluGlyIleValSerProAspThrCysGluGlyAsnLeu 816
1374 CCGTCACTCCATTCAGCGGGGCTACTGCATCCCAACAGGACATTCCTGGTGTTCATC 1433
817 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 836
1434 ACCTTCAAATATCAGCTTCACTACCTCCCTGAGACCTGGCTTCAATGCTGCTCAAGGAC 1493
837 LeuSerLeuSerIleGluArgGlyIleThrValAlaPheValGlySerSerGlyCysGly 856
1494 TTCACCTGAAGCTGCCCTCTGGCAAGATTGTGGCTCTGTGGGCCAGTCTGGGGGAGGA 1553
857 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 876
1554 AAGACCAAGTGTGCTCTGCTGGAGCGTCTTATGACCTGAAGCTGCTGCTGCTGCTGAGC 1613
877 PheAspGlyValAspAlaIleAlaIleGluAsnValGlnTrpLeuArgSerGln---IleAla 895
1614 TTGGATGGGATGACCTGCGAATCTCAACCTCTGCTGGCTCCGGGCGCAGGTATAGGT 1673
896 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 915
1674 TTCATCAGCAGGAGCCAGTCTCTGTTGCAACCAACCATCATGAGAGATATCCGATTGGG 1733
916 -----AspAsnSerArgValValProLeuAspGluIleIleLysGluAlaAlaAsnAlaAla 933
1734 AAGCTGATGCTCCGAT-----GAAGAGGTGTACACAGCTGCACGAGAGGCC 1781
934 AsnIleHisSerPheIleGluGlyLeuProGluIleIleValSerThrGlnValGlyLeuLys 953
1782 AATGCCACAGATTCATCAGCAGCTTCCCGATGGCTACAGCACTGTGGTTGGTGGACGG 1841
954 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 973

```

```

Db 1842 GGCACACACCTTCTGTGGTGGCCAGAGCAGCGCTAGCCATCGCAGCTGCCCTCATCAAG 1901
Qy 974 LysProIleIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 993
Db 1902 CAGCCACACAGTCTGATCCTGGACGAGGCCACAGTGGCTAGATCAGAAATCCGAGAGG 1961
Qy 994 ValValGlnHisAlaLeuAspIleAlaArgThrGlyArgThrCysLeuValValThrHis 1013
Db 1962 GTGTACAGAGGCGCTGACCGGCGCAGTGTGGTGGCGCAGCTGTGGTTCATTGGCCAC 2021
Qy 1014 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyIleLys 1033
Db 2022 CGCTTAGTACTTCTCGTGCAGCCCACTCCATCATGTGATGCCAATGCCAAGTCTGT 2081
Qy 1034 GluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 1053
Db 2082 GAGCTGGGACCCACGAGAACTCTTAAAGGGGGCTCTATTTCAGAGCTTATCCGG 2141
Qy 1054 AlaGlnSerVal 1057
Db 2142 AGACAAACCCGTG 2153

RESULT 10
BF584668 894 bp mRNA linear EST 12-DEC-2000
LOCUS 602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
DEFINITION mRNA sequence.
ACCESSION BF584668
VERSION BF584668.1 GI:11658386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9798 row: 1 column: 02
High quality sequence stop: 651.
Location/Qualifiers
FEATURES
source
1..894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4218385"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.6 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 263 a 203 c 242 g 185 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.33e-77 Length: 894
Score: 771.00 Matches: 171
Percent Similarity: 79.72% Conservative: 57
Best Local Similarity: 59.79% Mismatches: 48
Query Match: 14.47% Indels: 15
DB: 12 Gaps: 3

US-09-873-409-4 (1-1058) x BF584668 (1-894)

```

QY 129 ValPheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 148  
 DB 5 GTCTCTTTTCCGTTAAAT-GGAGCATTCAGTGTGGACAGGATCTCCAAATATTGAA 63  
 QY 149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 168  
 DB 64 GCCTTCGCCAATGACGAGGAGCAGCTTATGAAGCTTCAAAATATGATAATTAAGCCC 123  
 QY 169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188  
 DB 124 AGTATACAGAGCTTCTCAAGAGATGGGCACCAACACACACATACAAAGAAATCTCGAA 183  
 QY 189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208  
 DB 184 TTTAAGATATATCACTTCAGTACCTACCTCGAAAGAGTTCAGATCTTGAAGGGCCTC 243  
 QY 209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228  
 DB 244 AATCTGAAGGTGAAGAGCGGACAGCGGTGGCCCTGTGTGGCAACAGTGGCTGTGGAATA 303  
 QY 229 SerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248  
 DB 304 AGCAACAATGTCTCAGCTGATGCAAGGCTCTACAGCCCTTAGATGGCATGCTCAGTATC 363  
 QY 249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268  
 DB 364 GACGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGAGATCATGTGTGGTG 423  
 QY 269 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 288  
 DB 424 AGTCAGGAACCTGTGTGTGTGGCCACACGATCGCCGAGAACATTCGCTATGGCCGAA 483  
 QY 289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 308  
 DB 484 GATCTACCATGTGATGATGAGATTTGAAAGAGCTGTCAAGAGCCATGCTATGACTTCATC 543  
 QY 309 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 328  
 DB 544 ATGAACCTGCCCACTTTTGACCTGTGTGTGAGAGAGGGGCGAGCTGAGTGGG 603  
 QY 329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348  
 DB 604 GCACAGAAACAGAGAATCGC-ATTGCCGGGGCCCTGTGTCCGCAATCCAAAGTGCCTTC- 661  
 QY 349 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla-Le 368  
 DB 662 GTGACCAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGACCT 720  
 QY 368 uGluLysAlaSerLysGlyArgThrThrIle-----ValValAlaHisArgLe 384  
 DB 721 AGATAGGTACAAAGGCGGAGCACCATTTGGGAGCTCATCGCTGTGTACCGTCAAGCT 780  
 QY 384 uSerThrIleArgSerAlaAspLeuIleVal-----ThrLeuLysAspGlyMetLeuAl 402  
 DB 781 GACGGCCATGGCGGCGAAGGGGGG-CTTGTGGGAACGACTCTGAAAGC-----TC 830  
 QY 402 aGluLysGlyAlaHis 407  
 DB 831 TGAAGAGGTACTCAT 846  
 RESULT 11  
 BG248052 1019 bp mRNA linear EST 13-FEB-2001  
 LOCUS 602359987F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG248052  
 VERSION BG248052.1 GI:12757867  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1019)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 plate: LLAM10334 row: n column: 21  
 High quality sequence stop: 650.  
 FEATURES  
 source  
 1..1019  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4488404"  
 /clone\_lib="NCI\_CGAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 276 a 262 c 283 g 196 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,9e-77 Length: 1019  
 Score: 770.50 Matches: 154  
 Percent Similarity: 79.17% Conservative: 36  
 Best Local Similarity: 64.17% Mismatches: 47  
 Query Match: 14.46% Indels: 3  
 DB: 12 Gaps: 1  
 US-09-873-409-4 (1-1058) x BG248052 (1-1019)  
 QY 813 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 832  
 DB 4 GAAGGAATGTGCAATTTAGTGAGTGTGTTCACCTATCCACCCGACCCAGCATCCCA 63  
 QY 833 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 852  
 DB 64 GTGCTTCAGGGGCTGAGCCTTGAGTGAAGAGGGCCAGACGCTGGCCCTGGTGGGCGAGC 123  
 QY 853 SerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGln 872  
 DB 124 AGTGGCTGCGGGAGAGACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183  
 QY 873 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 892  
 DB 184 GGATCAGTGTCTTCTAGTGGCAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243  
 QY 893 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 912  
 DB 244 CAGCTGGGCAATTTGTCTCCCAAGAGCCCATCTCTTTTACTGTCAGCATCGCAGAGAACATT 303  
 QY 913 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 932  
 DB 304 GCCTACGGAGACACAGCCGGGTGGTCTTATGAGAGAGATTGTGAGGGCAGGCCAAGGAG 363  
 QY 933 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 952  
 DB 364 GCCAACATCCACCATTCATCGACTCGCTACCTGATAATAACACACAGTAGGAGAG 423  
 QY 953 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 972  
 DB 424 AAAGGCACTCAGCTGTGGGTGGGCGAAGAGCAGCGCATGCCATCGCAGCGCCCTCGTC 483

```

QY 973 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 992
    : : : : :
Db 484 AGACAGCCTCACATTTCCTGGACGAAGCAACATCAGCTCGGATACAGAAAGTGAA 543

QY 993 LysValValGlnHisAlaLeuAspLysAlaAaGThrGlyArgThrCysLeuValValThr 1012
    : : : : :
Db 544 AAGTTGTCCAGGAAGCGTGGCAAAAGCAGGAGGCGCACCTGCAATGTGATCGCT 603

QY 1013 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1032
    : : : : :
Db 604 CACCGTGTCCACCATCCAGAACGGGACTTGATCGTGGTGATTCAGAACCGGAAGG 663

QY 1032 leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
    : : : : :
Db 664 TCAAGGGAGCAGCGGACCCCAACAGCAAGTGTGGCGAGAGGCGATCTACTTC 719

RESULT 12
BG293345 726 bp mRNA linear EST 21-FEB-2001
LOCUS 602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
DEFINITION mRNA sequence.
ACCESSION BG293345
VERSION BG293345.1 GI:13052943
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 726)
JOURNAL NIH-MGC http://mgs.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10371 row: 1 column: 09
High quality sequence stop: 658.
FEATURES
    source
        1..726
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="IMAGE:4502552"
            /clone_lib="NIH MGC 94"
            /tissue_type="retina"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 3.3 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC_Library."
BASE COUNT 189 a 220 c 190 g 127 t
ORIGIN

Alignment Scores:
Pred. No.: 1,11e-75 Length: 726
Score: 753.00 Matches: 152
Percent Similarity: 84.29% Conservative: 25
Best Local Similarity: 72.38% Mismatches: 30
Query Match: 14.13% Indels: 3
DB: 12 Gaps: 0

US-09-873-409-4 (1-1058) x BG293345 (1-726)

QY 847 ValAlaPheValGlySerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArg 866
    : : : : :
Db 2 CTGGCCCTGTGGGACAGTGGCTGGCGGAAGACAGACAGTGGTCCAGCTGCTGAGCGC 61

```

```

QY 867 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlnLeuAsn 886
    : : : : :
Db 62 TTCTATGACCCCATGGCTGGATCAGTGTCTTTAGATGGTCAAGAAAGAACTCAAT 121

QY 887 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 906
    : : : : :
Db 122 GTCCAGTGGCTCCGAGTCACTGGGCATTGTGCCAGAAACCCATCTCTTTGACTGC 181

QY 907 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIle 926
    : : : : :
Db 182 AGCATCGCAGAGAACATCGCTATGGAGACAACAGCCGGTCTGTCCTCATGATGAGATT 241

QY 927 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 946
    : : : : :
Db 242 GTGAGGGCAGCCAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCAAAATAT 301

QY 947 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAla 966
    : : : : :
Db 302 AACAAAGAGTAGGAGACAAAGGGGACGACGCTCTCGGGGGCCAGAGCAGAGATTGCC 361

QY 967 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAla 986
    : : : : :
Db 362 ATGCCCGGAGCCCTCATCAGACAGCTCGGGTCTTACTCTGATGATGATGATGATGAT 421

QY 987 LeuAspAsnAspSerGlyLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1006
    : : : : :
Db 422 CTGGATCTAGAGTGAAGGTTGTCCAGGAAGCACTGCACAAAGCCAGGAGGCCGC 481

QY 1007 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 1026
    : : : : :
Db 482 ACCTGCATTTCGATCGCTCACCCTGTCACCATTCACAAACGCGGCACCTTGATCGTGT 541

QY 1026 lleuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 1046
    : : : : :
Db 542 GATTGGAACGCGCAAGGTCAAGGACGACGCGACCGCCACCAGCAGCTGCTGGCCGCAAGG 601

QY 1046 spileTyrPhe-LysLeuValAsn 1053
    : : : : :
Db 602 GCATCTATTTCCTCAACTGGTCAAC 625

RESULT 13
AZ682350 871 bp DNA linear GSS 14-DEC-2000
LOCUS ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION AZ682350
VERSION AZ682350.1 GI:11819496
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 871)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@igr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence stop: 16
High quality sequence start: 860.
High quality sequence start: 860.
Location/Qualifiers
    1..871
        /organism="Entamoeba histolytica"
FEATURES
    source

```



/strain="HMI:IMSS"  
/db xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/notes="Vector: pHS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

BASE COUNT 349 a 117 c 174 g 231 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.72e-75 Length: 871  
Score: 752.50 Matches: 151  
Percent Similarity: 71.64% Conservative: 46  
Best Local Similarity: 54.91% Mismatches: 75  
Query Match: 14.12% Indels: 3  
DB: 17 Gaps: 1

US-09-873-409-4 (1-1058) x AZ682350 (1-871)

Qy 153 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspIleAspAsn 172  
Db 45 GCTAAGTTCGCGCATTAAGTTATCAACAATTTGATAGATCCCGATATTGATTGT 104  
Qy 173 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 192  
Db 105 CAGTCTATTGGAGGTGAATGCCAAGTCACTGAGTGAATGAAATATTAGATTGGAGATGT 164  
Qy 193 SerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIle 212  
Db 165 CAATTTGTTTATCAACAGACTGTCTCATGATGATTAAAGACTTGACCTTGAATTT 224  
Qy 213 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 232  
Db 225 AAGAAAGGACAAACAATTTGATTTAGTTCGAGCATCAGGATGTGGAGTCAACTACTATT 284  
Qy 233 GlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAsp 252  
Db 285 CAATTAATCCAAAGAAATATTGATCCAAATGTGGAGAGTAACATTAGACCGGAAAGAT 344  
Qy 253 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 272  
Db 345 ATACGAGATTGATATCAATGTTAGAAATCAATAGATTAGTTGGACAGACCA 404  
Qy 273 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAsp----- 289  
Db 405 GTGTTGTTTCGAGGAACAATTCGAGAAATATTATGCTTCGAGCTAAAGAGGACCA 464  
Qy 290 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 309  
Db 465 CCAAGTGAAGAAGAGATGATTGAATGTCTAAATGGCAATTCACATGACTTCTTCT 524  
Qy 310 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGly 329  
Db 525 AAATCCGAGGAGGATATGACACAATAATTGGAGAAAGAGCAATTATTCAGAGGA 584  
Qy 330 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 349  
Db 585 CAAAACAAAGAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 644  
Qy 350 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGlu 369  
Db 645 GATGAAGCTACATCAGCATTCGATACACAAAGTGAAGAGATTGTACAAAGAGCATTTGA 704

Qy 370 LysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSer 389  
Db 705 AAAGCATCTAAAGGAAGAACAAATATTATTAGCACATAGACTAAACACTGTAGAAAT 764  
Qy 390 AlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlu 409  
Db 765 GCAGATAAAATATGTCTATTTCATCAAGAGAGAAATATTTCACACAGGAAACATCAAGAA 824  
Qy 410 LeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln 424  
Db 825 TTAATGATTTGAAGGAACATATTATGATTTAGTCAAGAGACAA 869  
RESULT 14  
BF796582  
LOCUS 602258463F1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4341710 5',  
DEFINITION mRNA sequence.  
ACCESSION BF796582  
VERSION BF796582.1 GI:12101636  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 944)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-romail.nih.gov  
Tissue Procurement: Louis Staudt, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM9955 row: n column: 15  
High quality sequence start: 9  
High quality sequence stop: 669.

FEATURES  
source

1. .944  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone\_lib="NIH MGC 85"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.867 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 265 a 241 c 237 g 201 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.94e-75 Length: 944  
Score: 751.00 Matches: 166  
Percent Similarity: 74.62% Conservative: 31  
Best Local Similarity: 62.88% Mismatches: 60  
Query Match: 14.09% Indels: 9  
DB: 12 Gaps: 2

US-09-873-409-4 (1-1058) x BF796582 (1-944)

Qy 779 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLysLysProAsn 798  
Db 39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAGACACCTCTG 98  
Qy 799 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 818  
Db 99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCCTGATAAATTTGAAGAAATATAACATTT 158

QY 819 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 838  
 DB 159 AATGAAGTCGTGTTCAACTATCCACCCGAGCAACGTCAGTCGTTCAGGGCTGAGC 218  
 QY 839 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 858  
 DB 219 CTGGAGTCGTAAGAAAGCCAGACACTAGCCCTGCTGGCGCAGCAGTGGCTGGGAAGAGC 278  
 QY 859 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 878  
 DB 279 ACGGTGTCAGCTCTGGAGCGGTTCTACGACCCCTTGGCGGGAGCAGTCTTCGAT 338  
 QY 879 GlyValAspAlaLysGluLeuAenValGlnTrpLeuArgSerGlnIleAlaIleValPro 898  
 DB 339 GGTCAAGAAAGCAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398  
 QY 899 GlnGluProValLeuPheAenCysSerIleAlaGluAenIleAlaTyrGlyAspAenSer 918  
 DB 399 CAGGAGCCTATCTCTATTGACTGAGCATTTGCCGAGAATATTGCCATGAGACAACAGC 458  
 QY 919 ArgValValProLeuAspGluIleLysGluAlaAlaAenAlaAenIleHisSerPhe 938  
 DB 459 CGGGTGTATCAGAGATGAATGCTGAGTGCAGCCAAAGCTGCCACATATCTCTTTC 518  
 QY 939 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 958  
 DB 519 ATCGAGACGTTATCCCCCAAAATATGAACAAAGAGTGGGAGATGAAGGAGCTCAGCTCTCA 578  
 QY 959 GlyGlyGlnLysGlnArgLeuAlaIleAlaAenAlaLeuLeuLys-ProLysIleLe 978  
 DB 579 GGAAGTGCACAAACAGAGGATGTGTTATGCCGAGCCCTCATCAGACAACCTCAAAATCT 638  
 QY 978 uLeuLeuAspGluAlaThrSerAla-LeuAspAenAspSerGluLysValVal- - - - -G 996  
 DB 639 CCGTGTGATGAAGCTATACATGAGCTCTGGATACCTGCAAGTGAACACCGTTGTCCAA 698  
 QY 996 IHisAlaLeuAspLysAlaArgThrGlyArgThrCysLysLeuValValThrHisArgLeuS 1016  
 DB 699 GAAGCCCTCTCACAGCAGAGAGCGCGCACCTGCTGTGATGCTCACCGCTGTC 758  
 QY 1016 exAlaIleGlnAenAlaAspLeuIleValValLeuHisAenGlyLysIleLysGluGlnG 1036  
 DB 759 AAC-ATCCAGAAATGCAGACTT-ATAGTGTGTTTCCCAACGGGAGAGTCAAG- - - - - 808  
 QY 1036 lyThrHis 1038  
 DB 809 --ACACAT 814

RESULT 15  
 BHI39685  
 LOCUS ENTN477R Entamoeba histolytica 886 bp DNA linear GSS 07-AUG-2001  
 DEFINITION Genomic, DNA sequence.  
 ACCESSION BHI39685  
 VERSION BHI39685.1 GI:15098746  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 886)  
 Loftus, B., Wang, Z., Van Aken, S., and Fraser, C.  
 AUTHORS Determination of clone end sequences from Entamoeba histolytica  
 TITLE HM1:IMSS sheared DNA library (2001)  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjloftus@igr.org  
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared

DNA library  
 Seq primer: M13-Reverse  
 Clans: shotgun  
 High quality sequence start: 18  
 High quality sequence stop: 851.  
 Location/Qualifiers  
 1. .886  
 /organism="Entamoeba histolytica"  
 /strain="HM1:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: PHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell. Oxford University Press, 1999)."  
 BASE COUNT 348 a 105 c 176 g 257 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1-93e-74 Length: 886  
 Score: 743.50 Matches: 151  
 Percent Similarity: 73.43% Conservatives: 48  
 Best Local Similarity: 55.72% Mismatches: 71  
 Query Match: 13.95% Indels: 13 2  
 DB: 17 Gaps: 1

US-09-873-409-4 (1-1058) x BHI39685 (1-886)

QY 786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805  
 DB 20 GCAAGAATACTTATGATGTTATTGATAGAAAACCACT-ATTGATTGTTATAGTGAAGAA 78  
 QY 806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825  
 DB 79 GGTGAACATTTTAATGATGTTAAAGGTGAATTTAAAGACATTTGTTTCAGATAT 138  
 QY 826 ProCysArgProAspValPheIleLeuArgGlyLysSerLeuSerIleGluArgGlyLys 845  
 DB 139 CCAACAGACCAACCAATCTGCTTGAAGGTATTTCATTCAAAGTAGAACAGGAAA 198  
 QY 846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865  
 DB 199 ACTGTTGCATTAGTAGGAGCATCAGGATGTTGTAATCAACATCAGTTCAGTTGATTGAA 258  
 QY 866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885  
 DB 259 AGATTTTATGATCCAAACACATCGAGATGTATTATTAGATGGACATAATATCAAGATTG 348  
 QY 886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905  
 DB 319 AATATTCATTTCTTAAGAGTCAATTTGGAATGATGTTAGGACAGACAGTATTATTGCT 378  
 QY 906 CysSerIleAlaGluAenIleAlaTyrGlyAspAenSerArgVal- - - ValProLeuAsp 924  
 DB 379 GAAAGTGTATGATAATATTAGAACGAGGAGTACCTAAAGGAGTTCAAGTAAGTAATGAA 438  
 QY 925 GluIleLysGluAlaAlaAenAlaAlaAenIleHisSerPheIleGluGlyLeuProGlu 944  
 DB 439 CAAATTTATGCTGCTGCTAAAATGGCAATGCACTGACTTTTATTTTCAGCAATGCCAGAA 498  
 QY 945 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArg 964  
 DB 499 GGATATAACACAATGGTAGGTGATAGAGGTGCACAAATTTTCAGGAGGACAAACAAAGA 558

```
Qy 965 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 984
Db 559 ATTGCTATTGCACGTGCATTGATTAGAAATCCAAAGTGTATTACTCGATGAAGCTACA 618
Qy 985 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1004
Db 619 TCACACTTGTATTCAGAAAGTGAAGAGATTGTACAAGATGCACCTTGACAAGCAGCAAAA 678
Qy 1005 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1024
Db 679 GGAAGACACACAATTGTAATTGCACATAGATTATCAACTATTCAAAATGCAGATCAATA 738
Qy 1025 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 1044
Db 739 TGTGTTATTATGACAGCAGAAGATTGCAGAAAGAGGACACATCAAGAGTTATTAGATTG 798
Qy 1045 ArgAspIleTyrPheLysLeuValAsnAlaGln 1055
Db 799 AAGGATTTTATTATACACTTGCTATGCACAA 831
```

Search completed: March 31, 2003, 13:50:31  
Job time : 7305.24 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 14303 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-4

Perfect score: 5330

Sequence: 1 MVISLTSKLSAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+p2n.model -DB=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-O=/cn2\_1/USPTO spoal/US09873409/runat\_27032003\_115418\_19228/app query fasta\_1.7544

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09873409 @CN 1 1.54887 @runat\_27032003\_115418\_19228 -NCPU=6 -ICPU=3

-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.ev.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.in.\*

18: em.mu.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.ste.\*

28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5328	100.0	3177	6	AX339030 Sequence
2	5328	100.0	3702	6	AX339031 Sequence
3	5159.5	96.8	3621	6	AX339032 Sequence
4	4985.5	93.5	3699	6	AX478104 Sequence
5	4561.5	85.6	2856	6	AX339028 Sequence
6	3456	64.8	2066	6	AX339027 Sequence
7	3137.5	58.9	4296	10	CRUGPI
8	3136.5	58.8	3987	10	CRUGPI1185
9	3136.5	58.8	4304	10	CRUGPI1165
10	3121.5	58.6	4018	5	GGA9799
11	3118.5	58.5	3858	4	GAU78609
12	3116.5	58.5	4927	10	AF257746
13	3115.5	58.5	4279	6	AX105082 Sequence
14	3112.5	58.4	4279	6	AX105057 Sequence
15	3112.5	58.4	4279	6	AX105078 Sequence
16	3112.5	58.4	4279	6	AX105080 Sequence
17	3111.5	58.4	4317	4	AF045016
18	3111.5	58.4	4317	6	AX105059 Sequence
19	3110.5	58.4	4045	12	AF269224
20	3105.5	58.3	3934	4	CFA413568
21	3103.5	58.2	4323	10	AF286167
22	3100.5	58.2	3860	6	AX322787 Sequence
23	3100.5	58.2	3860	6	AX322789 Sequence
24	3100.5	58.2	4186	6	AX108654 Sequence
25	3100.5	58.2	4195	6	AX108656 Sequence
26	3100.5	58.2	4264	6	AR051647 Sequence
27	3100.5	58.2	4264	6	AR051650 Sequence
28	3100.5	58.2	4646	6	AX336420 Sequence
29	3100.5	58.2	4646	6	AX336708 Sequence
30	3100.5	58.2	4646	6	AX391099 Sequence
31	3100.5	58.2	4646	6	I49610 Sequence 2
32	3100.5	58.2	4646	9	HUMDDR1
33	3100.5	58.2	4669	6	I08557 Sequence 3
34	3100.5	58.2	6505	6	AR028671 Sequence
35	3100.5	58.2	8630	6	AX012320 Sequence
36	3100.5	58.2	8630	6	AX012321 Sequence
37	3100.5	58.2	9318	6	AR028672 Sequence
38	3097.5	58.1	4378	6	E02326
39	3091	58.0	3888	10	AY082609
40	3090.5	58.0	3682	10	MUSPGLY1A
41	3090.5	58.0	4788	6	AX322793 Sequence
42	3090.5	58.0	4924	10	MUSMDR1A
43	3089	58.0	4280	10	CRUGPII
44	3088.5	57.9	4356	10	MUSMDRAA
45	3086	57.9	4233	6	AR123273 Sequence

# ALIGNMENTS

RESULT 1

AX339030  
LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 12 from Patent WO0194400.  
ACCESSION AX339030  
VERSION AX339030.1 GI:18129122  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
JOURNAL homologue on chromosome 7p15-21 and uses thereof  
Patent: WO 0194400-A 12 13-DEC-2001:  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
Location/Qualifiers  
source 1. 3177  
BASE COUNT 970 a 601 c 735 g 870 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3177  
Score: 5328.00 Matches: 1058  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.96% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-4 (1-1058) x AX339030 (1-3177)  
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20  
Db 1 ATGTCATCTCATTCAGCAGTAGGAATTAAGTCCTATTCCAAAGCTGGGGCTGTGGCA 60  
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40  
Db 61 GAAGAAGCTTGTTCATCAATCCGAACAGTCATAGCCCTTTAGGGCCCAAGGAAGAACTT 120  
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60  
Db 121 CAAAGGCTTTCTTTAAATATAACAAGATATGCTTGGTTTTATTTTCCCAAGTGGCTA 180  
Qy 61 LeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80  
Db 181 CTAAGTTGTGTCTCTTTTAAAGGTATACACAGAAATCTCAAGAGATGCCAAGGATTTT 240  
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100  
Db 241 GGCATATAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTACTTCTTTATGAAAT 300  
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120  
Db 301 GGAACCTATGACCTGCTTTTGGTATGGAACCTCTTGATTTCTTAATGAGAACCTGGA 360  
Qy 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140  
Db 361 TATACCATCGGAGCTGTCTTCTGCTTTCTTAGTGTAAATCCATAGCAGTATTGCAAT 420  
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
Db 421 GGACGAGAGCTCCCTCACTTTGAACACCTTCGCAATAGCCCCGAGAGCTGCCCTTCATAT 480  
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180  
Db 481 TTCCAGGTATTGATAAGAAACCCAGTATAGATAAATTTTCCACAGCTGGATATAAACCT 540  
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200  
Db 541 GAATCCATAGAAGAACTGTGGAATTTAAAAATGTTTCTTCAATATTATCCATCAAGACCA 600

201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
Db 601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCTTG 660  
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240  
Db 661 GTCGGTCTCAATGGCAGTGGAGAGTAGCGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720  
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260  
Db 721 CCGGATGATGGCTTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGCGCAT 780  
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280  
Db 781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTGGGACCAACCATCAGT 840  
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300  
Db 841 AACATATCAAGTATGACGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGG 900  
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
Db 901 GAAGCAAAATGCGPATGATTTTATCATGGAGTTTCCCTAATAAATTTAATACATTTGGTAGG 960  
Qy 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340  
Db 961 GAAAAAGGAGCTCAAAATGATGGAGGCGAGAAACACAGAGATCGCAATTTGCTCGCTTA 1020  
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
Db 1021 GTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTCGCCCTGGATTGAGAAAGC 1080  
Qy 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyValArgThrThrIleValVal 380  
Db 1081 AAGTCAGCTGTTCAAGCTGCACCTGGAGAGCGCAGAAAGGTGGAGTCAATTCGTGGTA 1140  
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
Db 1141 GCACACCGACTTTCTACTATTCCAGTGCAGATTTGATTGTGACCCCTAAAGGATGGAATG 1200  
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420  
Db 1201 CTGGCGGAGAAAGAGCAGACATGCTGAACCTAATGGCAAAACAGAGTCTATATTATTCATT 1260  
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
Db 1261 GTGATGTCAGAGATATTAAAAAGCTGTGAACAGATGGAGTCAATGACATATTCTACT 1320  
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460  
Db 1321 GAAAGAAAGACCAACTCACCTTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTTCATT 1380  
Qy 461 AspLysAlaGluGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeuLeu 480  
Db 1381 GACAAGCTCAGGAATCCACCAATCTAAAGAGATAGTCTTCTGAGGTCTCTCTATTA 1440  
Qy 481 LysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAlaSer 500  
Db 1441 AAAATTTTAAAGTTAAACAAGCTGAATGGCCTTTGTGGTTCTGGGACATTTGGCTTCT 1500  
Qy 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520  
Db 1501 GTTCTAAATGGAACGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTAATAACCATG 1560  
Qy 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540  
Db 1561 TTTGGAAATTAATGATAAAACACATTAAGCATGATGAGAAAATTTATTCATGATATTC 1620  
Qy 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560  
Db 1621 GTCATTTGGGTGTTATTTGCTTTGTGAGTTATTTATGCGGAGGATTTATTTACGGCAGA 1680  
Qy 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580

Db 1681 GCAGGGGAAATTTAAACGATGAGATTAAACACACTTGGCCCTTCAAAGCCCATGTTATATCAG 1740  
QY 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600  
Db 1741 GATATTGGCTGGTGGTGAAGGAAACAGACAGAGGGCTTGACAAATATTTAGCC 1800  
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620  
Db 1801 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAT 1860  
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640  
Db 1861 GCAACTTAACATGGGACTTTTCAGTTATCAATTCCTTTATATATATGATGGAGATGACATTC 1920  
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaIleMet 660  
Db 1921 CTGATTCGAGTATTGCTCCAGTACTTGGCCGACAGGAATGATTTGAACCCGAGCAATG 1980  
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680  
Db 1981 ACTGGATTTGCCAACAAAGATTAAGCAAGAACTTAAAGCATGCTGGAAGATAGCAACTGAA 2040  
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700  
Db 2041 GCTTTGGAGATATACGTACTATAGTGTCAATTAACAAAGGGAAGAAAGCCCTTCGAGCAATG 2100  
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720  
Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCAGATATTGGA 2160  
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740  
Db 2161 AGCTGTATGCATTACGCAATGCTTTATATATTTTGCCTATGCGACAGGGTTTCGATTT 2220  
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760  
Db 2221 CGAGCCTATTTAATTCAGCTGGACGAATGACCCAGAGGCGCATGTTTACTTTTACT 2280  
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780  
Db 2281 GCAATTTGCATATGGAGCTATGGCCATCGGAAACGCTCGTTTGGCTCTCGTAATATTC 2340  
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800  
Db 2341 AAAGCCAAATCGGGGCTGGGCATCTGTTGCTTGTGGAAAGAAACCAATATAGAC 2400  
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820  
Db 2401 AGCCGAGTCAGAGGGAAGAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCAGAA 2460  
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840  
Db 2461 GTCTCTTTCTTCATCCATGTCGCCCAGATGTTTTCATCTCCGTCGCTTATCCCTCAGT 2520  
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860  
Db 2521 ATTGACCGAGGAAGACAGTAGGATTTTGGGGAGCAGCGGCTGTGGAAAGCACTTCT 2580  
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyVal 880  
Db 2581 GTTCAACTTCGACAGACTTTATGATCCCGTCGCAAGGACAAGTGTGTTGATGGTG 2640  
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900  
Db 2641 GATGCAAAAGAAATGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAGAG 2700  
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920  
Db 2701 CCTGTGCTCTTCAACTGACGATTTGAGAAATCGCCTATGCTGATGACAAACGCCGTGTG 2760  
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940

Db 2761 GTGCATTAGATGAGATCAAAAGACCCGGAATGCGCAAAATATCCATCTCTTTATTGAA 2820  
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGly 960  
Db 2821 GGTCTCCCTCGAATAACAACACACAAGTTGGACTGAAAGGAGCAGAGCTTTCTGGCGC 2880  
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980  
Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTTATTGTTG 2940  
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000  
Db 2941 GATGAGCCACTTTCAGCCCTCGATTAATGACAGTGAAGGTGTTTCAGCATGCCCTTGT 3000  
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020  
Db 3001 AAAGCCAGGACGGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060  
QY 1021 AlaAspLeuIleValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040  
Db 3061 GCAGATTTGATAGTGGTTCGCACAATGGAAGATAAAGAAACAAGGAACCTCATCAAGAG 3120  
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058  
Db 3121 CTCCTGAGAAATCGAGACATATTTTAAATTAGTTAGTGAATGCACAGTCAGTGCAG 3174  
RESULT 2  
AX339031 3702 bp DNA linear PAT 09-JAN-2002  
LOCUS AX339031 Sequence 13 from Patent WO0194400.  
DEFINITION AX339031  
ACCESSION AX339031  
VERSION AX339031.1 GI:18129123  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1  
REFERENCE Frank,M.H. and Sayegh,M.H.  
AUTHORS A gene encoding a multidrug resistance human p-glycoprotein  
TITLE homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source  
1. 3702  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3702  
Score: 5328.00 Matches: 1058  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.96% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-4 (1-1058) x AX339031 (1-3702)  
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20  
Db 526 AUGTCTATCTCTGATTCACCATTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCA 585  
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeu 40  
Db 586 GAAGAAGTCTTGTCTCAATCCGAACAGTCATAGCCTTTAGGCGCCAGGAGAAAGAACTT 645  
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60  
Db 646 CAAAGGCTCTTCTTTAAATAATAACAAGATAGCTGTGTTTATTTCCTCCAGTGCCTA 705  
QY 61 LeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80

|||||  
706 CTAAGTTGTGTTCTGTTNTTGTAAAGGTATACACAGAATCTCAAGATGCAAAAGGATTTT 765  
QY GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsn 100  
Db GGCAATAAAGGACTATAGCTTCNAAGTGTCTCTGGTGTGTACTTCTTTATGAAT 825  
QY GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAenGlyGluProGly 120  
Db GGAACCTATGGACTTGTCTTTTGGTATGGAACTCTTGTATCTTAATGGAGAACCTGGA 895  
QY TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140  
Db TATACCATCGGACTGTCTTGTCTGTTTCTTTAGTGTAAATCCATAGCAGTTATTCGATT 945  
QY GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
Db GGAGCAGACGTCCCTCACTTTGAACCTTCGCAATAGCCGAGGAGCTGCTTTTCATATT 1005  
QY PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180  
Db TTCAGGTTATGTATAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAACCT 1065  
QY GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200  
Db GAATCCATAGAGGAACCTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 1125  
QY SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
Db TCTATCAAGATTCTGAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTG 1185  
QY ValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240  
Db GTCGGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1245  
QY ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAenValArgHis 260  
Db CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCAT 1305  
QY TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 280  
Db TATCGACCATATTTGAGTGGTGTAGTCAGAGGCTGTTTGTTCGGGACACCATCAGT 1365  
QY AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300  
Db AACATATCAAGTATGACGAGATGATGTGACTGATGAGAGATGGAGAGCAGCAGG 1425  
QY GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
Db GAAGCAAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAAATTAATCATTTGGTAGGG 1485  
QY GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340  
Db GAAAAAGGAGCTCAAAATGAGTGGAGGCGCAAAACAGAGGATCGCAATTTGCTGCGCTTA 1545  
QY ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
Db GTTCGAAACCCCAAGATCTTGATTTTATAGATGAGGCTACGCTCGCCCTGGATTCAGAAAGC 1605  
QY LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380  
Db AAGTCAGCTGTTCAGCTGCACTGGAGAGCGCAGCAAAAGTCGCACTACAATCGTGGTA 1665  
QY AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
Db GCACACGACTTCTACTATTCGAAGTGCAGATTGTGATTTGACCCCTAAAGGATGGAATG 1725  
QY LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420  
Db CTGGCGAGAAAGGAGACATGCTGAACTAATGGCAAAACGAGGCTCTATATTATTCATT 1785  
QY ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
|||||

Db 1786 GTGATGTCACAGGATATTAAAAAGCTGTATGAACAGATGGAGTCAATGACATATTCTTACT 1845  
QY GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460  
Db GAAAGAAAGACCACTCACTTCTCTGCACCTCTGTGAAGNCATCAAGTCAGACTTCATT 1905  
QY AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480  
Db GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCTCTGAAGTCTCTCTATTA 1965  
QY LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 500  
Db AAAATTTTAAAGTTAAACAAGCCCTGAATGGCTTTTGTGGTCTCTGGGACATTTGGCTTCT 2025  
QY ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520  
Db GTTCTTAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCCAAAATATTATACCATG 2085  
QY PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540  
Db TTTGGAAATAATGATAAAACCAACATTAAGCATGATGCAGAAATTTTATTCATGATATTTC 2145  
QY ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560  
Db GTCATTTTGGGTGTTATTTGCTTTGTTCAGTTATTTTCATGCAGGGATTTATTTACGGCAGA 2205  
QY AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580  
Db GCAGGGGAATTTTAACGATGAGATTAAAGACCTTGGCCCTTCAAGCCATGTTTATATTCAG 2265  
QY AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600  
Db GATATTTGCTGTTTGTATGAAAGGAAACACAGCACAGGAGGCTTGACAAATATTAGCC 2325  
QY IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620  
Db ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAAT 2385  
QY AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640  
Db GCAACTTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 2445  
QY LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660  
Db CTGATTTCTGATATTGCTCCAGTACTTCCGTGCACAGGAATGATTGAAACCGCAGCAATG 2505  
QY ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680  
Db ACTGGATTTGCCAACAAAGATAGCAAGAACTTAAAGCATGCTCGAAGAAAGATAGCAACTGAA 2565  
QY AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700  
Db GCTTTTGAGAAATATACGTACTATAGTGTCTATTAAAGGGGAAAAGCCCTTCGAGCAAAATG 2625  
QY TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720  
Db TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGACACAGATTATTGGA 2685  
QY SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740  
Db AGCTGTTATGCAATTCAGCCATGCTTTTATATATTTTGGCTTATGCACAGGGTTTCGATTTT 2745  
QY GlyValTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760  
Db GGAGCCTATTATTAATTCAGCTGGCAAGTAAGCCCAAGAGGCAATGTTTCATAGTTTTTACT 2805  
QY AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780  
Db GCAATTCGATATCGAGCTATGGCCTATGGCCATCGGAAAAAACCGCTCGTTTGGCTCTCGTAATTC 2865  
QY LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800  
Db AAAGCCAAATCGGGGCTCGCATCTGTTTGGCTTGTTCGAAAAAGAAACCAATATAGAC 2925



```
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2926 AGCGCAGTCAAGAGAGGAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAA 2985
QY 821 ValSerPhePheTyProCysArgProAspValPheIleuArgGlyLeuSer 840
Db 2986 GTCTCTTTCTTATCCATGTGCGCCAGAGTTTTCATCTCCGTGGCTTATCCCTCAGT 3045
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 3046 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGCTGTGGAAAGCACTTCT 3105
QY 861 ValGlnLeuLeuGlnArgLeuTyArgProValGlnGlnGlyValLeuPheAspGlyVal 880
Db 3106 GTTCAACTTCTCGAGAGACTTTATGACCCCGTCAAGGACAAAGTGTGTTCATGATGTG 3165
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 3166 GATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCTCAATAGCAATCGTTCCTCAAGAG 3225
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyArgAspAsnSerArgVal 920
Db 3226 CCTGTGCTCTCACTGCAGCATGTGTGAGACATCGCTATGTGACACAGCCGTGTG 3285
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
Db 3286 GTGCCATTAGATGAGATCAAGAAAGCGCAATGCGCAAAATATCCATTCTTTATTGAA 3345
QY 941 GlyLeuProGluTyArgThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
Db 3346 GGTCTCCCTGAGAAATACAAACACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGCGGC 3405
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeu 980
Db 3406 CAGAAACAAAGACTAGTATTGCAAGGCTCTTCTCCAAACCCAAATTTATTGTG 3465
QY 981 AspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3466 GATGAGGCCACTTCAGCCCTCGATATATGACAGTGAGAGGTGGTTCCAGCATGCCCTTGAT 3525
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3526 AAAGCCAGGACGGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3585
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3586 GCAGATTGATGTGTTCTGCACAAATGGAAGATGAAGAAACAGGAACACTCATCAAGAG 3645
QY 1041 LeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3646 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTCAATGCACAGTCAGTCAGCAG 3699

RESULT 3
LOCUS AX339032 3621 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 14 from Patent WO0194400.
ACCESSION AX339032
VERSION AX339032.1 GI:18129124
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 14 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source Location/Qualifiers
1..3621
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
BASE COUNT 1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 96.80% Indels: 27
DB: 6 Gaps: 1
US-09-873-409-4 (1-1058) x AX339032 (1-3621)
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAlaGlyAlaValAla 20
Db 526 ATGGTCACTCATTCACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGGCA 585
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeu 40
Db 586 GAAGAAGTCTTTGTCTCAATCCGAACAGTCATAGCTTTTAGGGCCCGAGGAAGAAGCTT 645
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPheProGlnTrpLeu 60
Db 646 CAA----- 648
QY 61 LeuSerCysValLeu***PheValArgTyThrGlnAsnLeuLysAspAlaLysPhe 80
Db 649 -----AGGTATACACAGAACTCTCAAGATGCAAGGATTTT 684
QY 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyPhePheMetAsn 100
Db 685 GGCATAAAAGAGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTACTTCTTTATGAT 744
QY 101 GlyThrTyGlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db 745 GAAACTATGACTTGTCTTTTGGTATGGAACCTCTTGATTTCTTAATGGAGAACCTGA 804
QY 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyCysIle 140
Db 805 TATACCATCGGAGCTTCTTGTCTTTCTTGTAGTAATCCATAGCAGTATTGTCATT 864
QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 865 CGAGCAGCAGTCCCTCCTCTTGTGAAACCTTCGCAATAGCCGAGGAGCTGCCTTTCATAT 924
QY 161 PheGlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyLysPro 180
Db 925 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 984
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyProSerArgPro 200
Db 985 GAATCCATAGAGGAACTGTGGAATTTAAATGTTCTTTCATATTCATCAAGACCA 1044
QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 1045 TCTATCAAGATTTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTG 1104
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArg 240
Db 1105 GTCCGTCTCAATGGCAGTGGGAAAGAGTAGCGTAGTCCAGCTTCTCGAGAGGTTATATGAT 1164
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 1165 CCGGATGATGGCTTTTATCATGTGTGATGAGATGACATCAGAGCTTTAAATGCGGCAT 1224
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 1225 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCCTGTTTGTTCGGGACCACCATCAGT 1284
QY 281 AsnAsnIleLysTyGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db 1285 AACAAATATCAAGTATGGACGAGATGATGTGCTGATGAAGAGATGGAGAGAGCAAGG 1344
```

QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
DB 1345 GAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAATTTAATACATGGTAGGG 1404  
QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340  
DB 1405 GAAAAAGGAGCTCAAAATGAGTGGAGGCGCAAAACAGAGGATCGCAATTTGCTCGTGGCTTA 1464  
QY 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
DB 1465 GTTCGAACCCCAAGATTTCTGATTTTAGAGCTACGGCTTCGCCCTGGATTCAGAAAGC 1524  
QY 361 LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380  
DB 1525 AAGTCAGCTGTTCAAGCTGCATCTGGAGAGGCGAGCAAGGTCCGACTACATCTGGTGA 1584  
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
DB 1585 GCACACCGACTTTCTACTATTTCGAAGTCAGATTTTGATTGTGACCCCTAAAGGATGGAATG 1644  
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420  
DB 1645 CTGGCGGAGAAAGGAGCACATGCTGAACCTAATGGCAAAACGAGGTCTATATTATTACATT 1704  
QY 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
DB 1705 GTGATGTCACAGATATTAAAAAGCTGATGAACAGATGGAGTCATGACATATTCTACT 1764  
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460  
DB 1765 GAAAGAAAGACCAACTCACCTTCCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCAATT 1824  
QY 461 AspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGluValSerLeuLeu 480  
DB 1825 GACAAGGCTCAGGAATCCACCAATCTAAAGAGATAGTCTTCCCTGAAGTCTCTCTATTA 1884  
QY 481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500  
DB 1885 AAAATTTTAAAGTTAAACAAGCTGAAATGGCTTTTGTGGTTCTGGGGACATTTGGCTTCT 1944  
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520  
DB 1945 GTTCTAAATGGAACTGTTTATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACCATG 2004  
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540  
DB 2005 TTTGGAAATTAATGATAAAACCAATTAAGCATGATGACAGAAATTTATTCCATGATATTC 2064  
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560  
DB 2065 GTCAATTTGGGTGTTATTGCTTTGTGTCAGTTATTTTCATGACAGGATTTATTTACGGCAGA 2124  
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580  
DB 2125 CGAGGGGAAATTTTAACGATGAGATTAAAGACACTTGGCCCTTCAAGCCATGTTATATACG 2184  
QY 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600  
DB 2185 GATATTCCCTGGTGTGATGAAAGGAAAAACAGCAGAGGCTTGCAACAAATATTAGCC 2244  
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620  
DB 2245 ATAGATATAGCACAAAATTAAGGAGCAACAGGTTCCAGGATTTGGCTCTTAAACAAATAT 2304  
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640  
DB 2305 GCAACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGATGGAGATGACATTC 2364  
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660  
DB 2365 CTGATTTCTGATATTGCTCCAGTACTTTGCCGTGACAGGAATGATTGAACCGCAGCAATG 2424

QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680  
DB 2425 ACTGGATTTGCCAAACAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2484  
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700  
DB 2485 GCTTTGGAGAAATATACGTACTATAGTGTCTTAAACAAGGGAAAAAGCCCTTCGAGCAATG 2544  
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720  
DB 2545 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATGGA 2604  
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740  
DB 2605 AGCTGTTATGCACTTCAGCCATGCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2664  
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760  
DB 2665 GGAGGCTATTTAAATCAAGCTCGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTTACT 2724  
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780  
DB 2725 GCATTTGCATATGGAGCTATGGCCATCGGAAAAACGCTGTTTGGCTCTCGAATATTC 2784  
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800  
DB 2785 AAAGCCAAATCGGGGCTGCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAATATAGAC 2844  
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820  
DB 2845 AGCCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCAGAA 2904  
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840  
DB 2905 GTCTCTTTCTTATCCATGTGCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 2964  
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860  
DB 2965 ATTGACGAGGAAAGACACAGTACGATTTTGTGGGAGCAGCGGCTGTGGGAAAGACATTCT 3024  
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880  
DB 3025 GTTCAACTTCTCAGAGACTTTATGACCCCGTGCAGAGCAAGTGTCTTGTATGTTGTG 3084  
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900  
DB 3085 GATGCAAAAGAAATTAATGATGACGTGCTCGCTTCCCAATAGCAATCGTCTCTCAAGAG 3144  
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920  
DB 3145 CCTGTGCTCTTCAACTGACGATTTGCTGAGAACATCGCCTATGTTGGTGACACAGCCGTGTG 3204  
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940  
DB 3205 GTGCCATTAGATGAGATCAAGAAGCCGAAATGCAAGCAAAATATCCATCTCTTTATTGAA 3264  
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyValAlaGlnLeuSerGlyGly 960  
DB 3265 GGTCCTCCCTGAGAAATACAACACAGATTTGGACTGAAGGAGCACAGCTTCTGCGCGC 3324  
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980  
DB 3325 CAGAAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAAACCCAAAAATTTTATTGTTG 3384  
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000  
DB 3385 GATGAGGCCACTTCAGCCCTCGATAATGACGTGAGAAAGTGTTCAGCATGCCCTTGAT 3444  
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020  
DB 3445 AAAGCCAGGACGGGAGGACATGCTTAGTGGTCACACAGGCTCTCTGCAATTCAGAAC 3504  
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040

```
Db 3505 GCAGATTGATAGTGGTCTGCGCAATAGGAAAGATAAAGGAAACAAGAACTCATCAAGAG 3564
|||||
Qy 1041 LeuLeuArgAsnArgAspIleTyrPheLeuValAsnAlaGlnSerValGln 1058
|||||
Db 3565 CTCTCGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTCGAG 3618
|||||
RESULT 4
AX478104
LOCUS
DEFINITION
Sequence 34 from Patent WO240541.
ACCESSION
AX478104
VERSION
AX478104.1 GI:22217064
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, X.,
Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,
Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G.,
Azinzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,
Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,
Policky, J.B. and Kearney, L.
Transmitters and ion channels
Patent: WO 0240541-A 34 23-MAY-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..3699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7472030CB1"
BASE COUNT 1116 a 707 c 860 g 1016 t
ORIGIN
Alignment Scores:
Pred. No.: 2e-317 Length: 3699
Score: 4985.50 Matches: 1002
Percent Similarity: 94.80% Conservative: 1
Best Local Similarity: 94.71% Mismatches: 2
Query Match: 93.54% Indels: 53
DB: 6 Gaps: 2
US-09-873-409-4 (1-1058) x AX478104 (1-3699)
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
|||||
Db 682 ATGGTCATCTCATGTGACCAAGTAAGGAATTAAGTGCCTTAITCCAAAGCTGGGGCTGTGGCA 741
|||||
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
|||||
Db 742 GAGAAGTCTTGTTCATCAATCCGACAGTCATAGCCCTTAGGGCCCGACGAGGAAGAACTT 801
|||||
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
|||||
Db 802 CAA----- 804
Qy 61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
|||||
Db 805 -----AGGTATACACAGAAATCTCAAGATGCAAAAGGATTTT 840
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
|||||
Db 841 GGCATATAAAGGACTATAGCTTCAAAAGTGCTCTTGGTGCTGTGACTTCTTTATGAAT 900
|||||
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
|||||
Db 901 GGAACCTATGGACTTGCTTTTGGTATGGAACCTCCTTGATTCTTAATGAGAACCTGGA 960
|||||
Qy 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
|||||
Db 961 TATACCATCGGACTGTCTTCTTGCTTTTCTTTAGTGTATATCCATAGTAGTATTGCAATT 1020
|||||
```

```
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
|||||
Db 1021 GGAGCAGCAGTCCCTCCTTTCACCTTTGAAACCTTCGCAATAGCCGAGAGCTGCTTTCATATT 1080
|||||
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
|||||
Db 1081 TTCCAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCT 1140
|||||
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
|||||
Db 1141 GAATCCATAGAAGAACTGTGGAATTTAAATAATGTTCTTTCATATTATCCATCAAGGCCA 1200
|||||
Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
|||||
Db 1201 TCTATCAAGATTCTGAAGGCTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCGCTTG 1260
|||||
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAsp 240
|||||
Db 1261 GTCGGTCTCAATGGCAGTGGGAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1320
|||||
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
|||||
Db 1321 CCGGATGATGGCTTTTATCATGTGTGATGAGATGACATCAGAGCTTTAAATGTGCGCAT 1380
|||||
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
|||||
Db 1381 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCACT 1440
|||||
Qy 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
|||||
Db 1441 AACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAG 1500
|||||
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
|||||
Db 1501 GAAGCAATCGGTATGATTTTATCATGGAGTTTCCATATAATTTAATACATTTGGTAGG 1560
|||||
Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
|||||
Db 1561 GAAAGAGAGCTCAAATGATGGAGGCGAGAAACAGAGATCGCAATTGCTCGTGCTTA 1620
|||||
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
|||||
Db 1621 GTTCGAAACCCCAAGATTCTGATTTAGTAGGCTACCTCTGCCCCCTGGATTGAGAAAGC 1680
|||||
Qy 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
|||||
Db 1681 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGCGCAGCAAGGTCGGACTACAAATCGTGTA 1740
|||||
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
|||||
Db 1741 GCACACCGACTTCTTACTATTTCGAAGTCAGATTTGATTGTGACCCCTAAAGGATGGAATG 1800
|||||
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
|||||
Db 1801 CTGGCGGAGAAAGGACGCGCATGCTGAACCTAATCGGAAAAACGAGGTCTATATTATTCACTT 1860
|||||
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
|||||
Db 1861 GTGATGTCAAGGATATTAATAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1920
|||||
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
|||||
Db 1921 GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1980
|||||
Qy 461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
|||||
Db 1981 GACAAGGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTATTA 2040
|||||
Qy 481 LysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAlaSer 500
|||||
Db 2041 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCCTTTTGTGGTCTTCTGGGGACATTTGGCTTCT 2100
|||||
```

```
QY 501 ValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThrMet 520
Db 2101 GTTCTAAATGGAACGTGTTTCATCCAGTATTTTCCATCATCTTTCCAAAATAATTAACCATG 2160
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleLysSerMetIlePhe 540
Db 2161 TTTTGGAAATTAATGATAAACCACATTAAGCATGATGACAGAAATTTATTCCATGATATTC 2220
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db 2221 GTCAATTTGGGTGTTATTTCTGTTGTCAGTTATTTTCAGT----- 2259
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
Db 2260 -----CAG 2262
QY 581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db 2263 GATATTCCTCGTTTGTATGAAAGGAAACAGCACAGAGGCTTGCAACAATATTAGCC 2322
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db 2323 ATGATATAGACAAAATTCAGAGAGCAACAGGTTCCAGGATTCGCGTCITTAACACAAAAT 2382
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640
Db 2383 GCAACTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGAGATGACATTC 2442
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db 2443 CTGATTCTGAGTATTCCTCAGTACTTGGCGTGACAGGAATGATTGAAACCGCAGCAATG 2502
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db 2503 ACTGGATTTGCCAACAAAGATAAGCAGAACTTAAGCATGCTGGAAGATAGCAACTGAA 2562
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db 2563 GCTTTGGAGAAATATACGTACTATAGTGTCTTAACAAGGGGAAAGCCTTCGAGCAATG 2622
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db 2623 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTTATGGA 2682
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db 2683 AGCTGTTATGCATTCAGCCATGCTCTTATATATTTTCCTATGCGGAGGTTTCGATTT 2742
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db 2743 GGAGCCTATTAAATCAAGCTGGACGAAATGACCCAGAGGGCATGTTTCATAGTTTACT 2802
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db 2803 GCAATTCATATGAGCTATGGCCATCGGAAACCGCTGCTTTTGGCTCTCGAATATTC 2862
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLysLysProAsnIleAsp 800
Db 2863 AAAGCCAAATCGGGGCTGCGCATCTGTTTTCCTTTGTTGGAAGAAACCAATATAGAC 2922
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
Db 2923 AGCCGCGAGTCAAGAGGGAAGAAACCCAGACACATGTGAAGGAATTTAGAGTTTCGAGAA 2982
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
Db 2983 GTCCTTTCTTCTATCCATGTCGCCACAGATGTTTTCATCTCCCTCGTGGCTTATCCCTCAGT 3042
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
Db 3043 ATTGAGCGAGGAAAGACAGTATGCAATTTTGGGGAGCAGCGCTGTGGGAAAGCACTTCT 3102
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
```

```
Db 3103 GTTCAACTTCTCAGAGACTTTATGACCCCGTGCAGGAGCAAGTCTGTTTATGCTGTG 3162
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
Db 3163 GATGCAAAAGAAATGATGATGACATGGCTCCGTTCCCAANTAGCAATCGTTCTCAAGAG 3222
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
Db 3223 CCTGTGCTCTTCAACTGCAGCATTCGTGAGAACATCGCCTATATGTTGACAAACCGCTGTG 3282
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
Db 3283 GTGCCATTAGATGAGATCAAGAAGCCGCAATTCGAGCAAAATATCCATTCTTTATTGAA 3342
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGly 960
Db 3343 GGTCTCCCTCGAGAAATACAAACACACAAAGTTGGACTGAAGGAGCACAGCTTTCTGGCGGC 3402
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 980
Db 3403 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAAATTTATTGTTG 3462
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3463 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAAGAGTGGTTTCAGCATGCCTTGAT 3522
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3523 AAAGCCAGAGCGGGAAGGACATGCCCTAGTGGTCACTCAAGGCTCTCTCAATTCAGAAC 3582
QY 1021 AlaAspLeuIleValIleLysHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3583 GCAGATTGATAGTGTCTTCGCAATGGAAGATAAAGAACAGAACTCATCAAGAG 3642
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3643 CTCCTGAGAAATCGACATATATTTAAAGTTAGTGAATGCACAGTCAGTCAGTCAG 3696
RESULT 5
AX339028 2856 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 10 from Patent WO0194400.
DEFINITION AX339028
ACCESSION AX339028
VERSION AX339028.1 GI:18129120
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank, M.H. and Sayegh, M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
JOURNAL homologue on chromosome 7p15-21 and uses thereof
PATENT: WO 0194400-A 10 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
Location/Qualifiers
1..2856
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 888 a 540 c 652 g 776 t
ORIGIN
Alignment Scores:
Pred. No.: 7,66e-290 Length: 2856
Score: 4561.50 Matches: 912
Percent Similarity: 97.96% Conservative: 0
Best Local Similarity: 97.96% Mismatches: 1
Query Match: 85.58% Indels: 18
DB: 6 Gaps: 1
US-09-873-409-4 (1-1058) x AX339028 (1-2856)
```

QY 129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGI 148  
DB 114 GTTTCCTTTTAGTGTAATCCATAGCAGTTATGTCATTTGGAGCAGCAGTCCTCATTTAT -- 171  
QY 148 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 168  
DB 172 -----GATAAGAAACC 182  
QY 168 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGI 188  
DB 183 CAGTATAGATAACTTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGTGA 242  
QY 188 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 208  
DB 243 ATTTAAATAATGTTCTTTCAATTTATCATCAAGACCATCTATCAAGATTTCTGAAGGTCT 302  
QY 208 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 228  
DB 303 GAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTTTGGTGGTCTCAATGGCAGTGGAA 362  
QY 228 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 248  
DB 363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGTG 422  
QY 248 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 268  
DB 423 GGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCATTTATCCAGACCATATTTGGAGTGT 482  
QY 268 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgSe 288  
DB 483 TAGTCAAGAGCCTTTGTTTGGGACCCACCATCAGTAACAATATCAAGTATGGAGAGA 542  
QY 288 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheI 308  
DB 543 TGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGAAGCAAAATGCGTATGATTTAT 602  
QY 308 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGI 328  
DB 603 CATGGAGTTTCTTAATAAATTTAATACATTTGGTAGGGGAAAAGGAGCTCAAAATGAGTGG 662  
QY 328 vGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuI 348  
DB 663 AGGGCAAGAACAGAGATCGCAATTCCTGTCCTTAGTTTCGAACCCCAAGATTTCTGAT 722  
QY 348 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaI 368  
DB 723 TTTAGTAGAGGTACGCTGCGCTGATTCAGAAACCAAGTCAGCTGTTCAAGCTGCAC 782  
QY 368 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 388  
DB 783 GGAGAAAGCGCAGCAAAAGGTCCGACTACAATCGTGGTAGCACACCGACTTTCTACTATTGG 842  
QY 388 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 408  
DB 843 AAGTGCAGATTTGATTTGACCTTAAGGATGAATGCTGGCGGAGAAAGGAGCAGCATGC 902  
QY 408 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 428  
DB 903 TGAACCTAATGGCAAAACAGGCTCTATATTATTCATCTTGATGATGCACAGATATTAAAA 962  
QY 428 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 448  
DB 963 AGCTGATGAACAGATGGAGTCAATGACATATCTACTGAAAGAACCAACCACTTCACCTCC 1022  
QY 448 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGI 468  
DB 1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTTGACAGGCTGGAGGAATCCACCCA 1082  
QY 468 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 488  
DB 1083 ATCTAAAGAGATTAAGTCTCTCTGGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACCAAGCC 1142  
QY 488 oGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 508

DB 1143 TGAATGGCCCTTTTGGTCTCTGGGACATTTGGCTTCTGTCTTAAATGGAACTGTTTCATCC 1202  
QY 508 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 528  
DB 1203 AGTATTTTCCATCATCTTTGCAAAATTTATAACCATGTTTGGAAATTAATGATAAACAC 1262  
QY 528 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 548  
DB 1263 ATTTAAAGCATGATGCAGAAATTTATTCATGATATTTCGTCAATTTGGTGGTATTATTGCTT 1322  
QY 548 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 568  
DB 1323 TGTCTAGTTATTTTCATGCAGGGATTTTACGGCAGAGCAGGGGAAATTTTAAACGATGAG 1382  
QY 568 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlyLy 588  
DB 1383 ATTTAGACACTTGGCTTCAAGCCATGTTATATCAGGNATTTGCTGGTGGTGGATGAAA 1442  
QY 588 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGI 608  
DB 1443 CGAAACACAGCACAGGAGGCTTCACAACAATATTAGCCATAGATATAGCAAAATTCAGG 1502  
QY 608 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 628  
DB 1503 AGCAACAGGTTCCAGGATTTGGCTCTTAACACAAATGCAACTAACATGGGACTTTTCAGT 1562  
QY 628 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 648  
DB 1563 TATCATTTCTTTATATATGATGGATGGAGATGACATTTCTGATTTCTGATTTGCTCCAGT 1622  
QY 648 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 668  
DB 1623 ACTTGGCGTGACAGCAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682  
QY 668 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 688  
DB 1683 GCAGAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTTGGAGAAATATACGTACTAT 1742  
QY 688 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGI 708  
DB 1743 AGTGTCATTTAACAGGGAAAAAGCCTTCGAGCAATGATGATGAGAGATGCTTCAGACTCA 1802  
QY 708 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 728  
DB 1803 ACACAGAAATACCTCGAAGAAAGCAGACAGATTTTGGAAAGCTGTTATGATTCAGCCATGC 1862  
QY 728 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGI 748  
DB 1863 CTTTATATATTTTGGCTATTCGACGAGGGTTTCGATTTTGGAGCTTATTTAAATTCAGGCTGG 1922  
QY 748 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 768  
DB 1923 ACGAATGACCCAGAGGGCATGTTTCATAGTTTTTACTGCATTTGCATATGGAGCTATGSC 1982  
QY 768 alIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 788  
DB 1983 CATCGMAAAACGCTCGTTCCTGCTGCTGAATATTCCAAAGCCAAATCGCGGCTGCGCA 2042  
QY 788 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 808  
DB 2043 TCTGTTTGGCTTTGTTGGAAAAAGAAACCAAAATATAGACGCCGCGCATGCTCAAGAGGGGAAAAA 2102  
QY 808 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 828  
DB 2103 GCCAGACACATGTGAAGGAAATTTAGAGTTTCGAGAAGTCTCTTCTCTATCCATGTCG 2162  
QY 828 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 848  
DB 2163 CCCAGATGTTTTCATCCTCCGTTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222  
QY 848 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 868





Db	1023	ACCTCGAAGAACACAGATTATTGGAAGCTGTATTGCAATTCAGCCATGCTTTATATAT	1082
Qy	732	PheAlaTyxAlaAlaGlyPheArgPheGlyAlaTyxLeuIleGlnAlaGlyArgMetThr	751
Db	1083	TTTGCCATGACGAGCGTTTCGATTGGAGCCTATTTAATTCAGCTGCAGAAATGACC	1142
Qy	752	ProGluGlyMetPheIleValPheThrAlaIleAlaTyxGlyAlaMetAlaIleGlyLys	771
Db	1143	CCAGAGGGCATGTTTCATAGTTTTTACTGCAATTCATATGGAGCTATGGCCATCGGAAA	1202
Qy	772	ThrLeuValLeuAlaProGluTyxSerLysAlaLysSerGlyValAlaHisLeuPheAla	791
Db	1203	ACGCTCGTTTGGCTCTCAATATTCAGAGCCAAATCGGGGCTCGCATCTGTTTGGC	1262
Qy	792	LeuLeuGluLysLysProAsnLeuAspSerArgSerGlnGluGlyLysLysProAspThr	811
Db	1263	TTGTTGAAAGAAACCAATATAGACAGCGCAGTCAAGAGGAAAGAGCCAGACACA	1322
Qy	812	CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyxProCysArgProAspVal	831
Db	1323	TGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTCTTCTATCCATGTCGCCAGATGT	1382
Qy	832	PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly	851
Db	1383	TTCATCTCCGTGGCTTATCCCTCAGTATTGACGAGGAAAGACAGTAGCATTTGTGGGG	1442
Qy	852	SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyxAspProVal	871
Db	1443	AGCAGCGCTGTGGGAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCGTG	1502
Qy	872	GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg	891
Db	1503	CAAGGCAAGTGTCTGTTGATGTTGGATGCAAAAGAAATTTGAATACAGTGGCTCCGT	1562
Qy	892	SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn	911
Db	1563	TCCCAATAGCAATCTGTTCTTCAAGAGCTGTGCTCTTCAACTGCGAGCTGCTGAGAAC	1622
Qy	912	IleAlaTyxGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn	931
Db	1623	ATCGCTATGTTGACAAACAGCGCTGTGTTGCCATTAGATGAGATCAAGAGCCGCAAT	1682
Qy	932	AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyxAsnThrGlnValGly	951
Db	1683	GCAGCAAAATATCCATTTCTTTATTGAAGTCTCCCTGAGAAATACAAACACAAGTTGGA	1742
Qy	952	LeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu	971
Db	1743	CTGAAAGGAGCACAGCTTTCTGGCGGCCGAAACAAAGACTAGCTATTGCAAGGGCTCTT	1802
Qy	972	LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer	991
Db	1803	CTCCAAAACCCCAATTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT	1862
Qy	992	GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal	1011
Db	1863	GAGAAGTGTTCAGCATGCTTGTATAAGCCAGACGCGAAGGACATGCTAGTGGTC	1922
Qy	1012	ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys	1031
Db	1923	ACTCACAGGCTCTCTCAATTCAGAACGAGATTTGATAGTGGTTCTGCACATGGAAG	1982
Qy	1032	IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyxPheLysLeu	1051
Db	1983	ATTAAGGAAACAGGAACACTCATCAAGAGCTCTCTGAGAAATCGAGACATATATTTAAGTTA	2042
Qy	1052	ValAsnAlaGlnSerValGln	1058
Db	2043	GTGAATGCACAGTCAGTGCAG	2063
RESULT 7			
CRUFGPI			
LOCUS			

DEFINITION	C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.		
ACCESSION	M60040		
VERSION	GI:191164		
KEYWORDS	multidrug resistance; p-glycoprotein; transmembrane protein.		
SOURCE	C.griseus adult liver and ovary, cDNA to mRNA.		
ORGANISM	Cricetulus griseus		
REFERENCE	1 (bases 1 to 4296)		
AUTHORS	Endicott, J.A., Sarangi, F. and Ling, V.		
TITLE	Complete cDNA sequences encoding the Chinese hamster P-glycoprotein		
JOURNAL	GENE		
MEDLINE	DNA Seq. 2 (2), 89-101 (1991)		
PUBMED	92135896		
FEATURES	1685679		
source			
Location/Qualifiers			
1..4296			
/organism="Cricetulus griseus"			
/db_xref="taxon:10029"			
/tissue_type="liver and ovary"			
/dev_stage="adult"			
103..3933			
/note="putative"			
/codon_start=1			
/product="p-glycoprotein isoform I"			
/protein_id="AAA68883.1"			
/db_xref="GI:191165"			
/translation="MEFEEDFSGRKDNFLKMGKSKKEKKEKPVVSVFTMRYAGH			
LDRLVGLTUALIHGVALFLMFLVDFGMDTDSFASVGNIPITNATNATQVNSADIFG			
KLEEMTYAYTIGAGVLIVAYIQVSWFLAAGRQIHKIRKQKFFHAINQWQIEGW			
DVHDVGLNLTLDVSKINGEIVGKIGMFMATFFGFIIGTRGKMLTLVLAI			
SPVLSGAGIWAHLSFTDLQKAKAGAVEVLAIRTVAFGQKGLKELRYNN			
NLEAKRLGIKKAITNSMGAAFLIIVASVALAFWYCTSLVSKYSGIOLVTVFA			
VLIAGFSGOASPNIEAFANARGAAVEIFNLIIDNKPSIDSPKNGYKPDNKGLEFK			
NHYSFSRQVQLKGLNKVQSGQTVLVNCGSGKSTTVQLLQRLVDFEGVWSI			
DQDIRITNRYLREIIGVVSQEPVLPATTAENIRYGRNVMTDEIEKAVEANAYD			
FIMLPHKFDLTGVRGAQSGQKQRIARALVRNPKILLDEATSDALDTESEAVV			
QALDKARQRTTIVIAHRLSTVRNADIAGFDGVIIVEQGNHLEMKRKGIFYKLVM			
TOTAGNEIELNVEGSKNEIDNLDMSKDSASSLIRRRSTRSRIRGPHDQDKLSLK			
EALDEDVPPISFWRLKLNSEMPYFVVGIFCALVNGALQPAFSIIISKVGVFTNT			
DDETKRDSNLSFLILGIVSIFITFLOQFTFGKAGEIITKLRYWVFKSMLRQDV			
SWFNPKNTGALTATRLANDAGQVKGATGARLAVITQNIANGTIGIILISLWQUTL			
LLLAIVPIITAGVVMKLSQALDKKLEGGSKIAATEIENFTVTVSTREQKFE			
NVYASQIQIPYRNALKKAHVFGITFTQAMMYFSAACFRFAGLVARELMTFENVL			
LVSALVFGMAVGVSSFADYAKAKVSASHIIMIIEKVPISIDSYSTGGLKPNTEG			
NVKNVEVNYPTRPDIPLVQLNLEVKKGTALVSGSGCGSKTVVOLLERFYDPMA			
GTVFDGKEVNLNVQMLRAHLGIVSQBPILFDCSIAENIAYGNSRVVSQDEIERAA			
KEANIHOFIESLPDKYTRVGDKGTQLSGGQKQRIARALVRQPHILLDEATSDALD			
ITYFSWVSVOAGAKR"			
BASE COUNT	1271 a	892 c	1029 g 1104 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2,61e-196	Length:	4296
Score:	3137.50	Matches:	598
Percent Similarity:	75.82%	Conservative:	211
Best Local Similarity:	56.04%	Mismatches:	215
Query Match:	58.86%	Indels:	43
DB:	10	Gaps:	6
US-09-873-409-4 (1-1058) x CRUGPI (1-4296)			
Qy	4	SerLeuThrSerLysGluLeuSerAlaTyxSerLysAlaGlyAlaValAlaGluVal	23
Db	805	TCATTACTGATAAAGAACTTCAGGCATATGCAGAGCTGCTGCTGAGAGTGC	864
Qy	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
Db	865	TTAGAGCCCATCAGAACTGTGATTCGATTTGGAGGACAAAGAAAGAACTTGA	918
Qy	44	PheLeuLeuAsnIleThrArgTyxAlaTrpPheTyxPheProGlnTrpLeuLeuSerCys	63



```
Db 918 ----- 918
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 919 -----AGGTCAACAACAATTTAGAACAGCCAAAGACATTGGGATAAG 963
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 964 AAGCGATCACAGCCACATTTTCATGGCGCAGCTTTCCTTATCATCATCATAT 1023
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1024 GCTCTGGCTTCTGGTATGGAACTCTCTGGTATCTCAAAAGAA-----TATTTATT 1077
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1078 GGCAAGTGCTCACTGCTCTTCTTGGTGTATTAATGGGCGATTCAGTATTTGGACAGCA 1137
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1138 TCTCCAATATCGAAGCTTTGCCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1198 ATTGATAATAAGCCCACTATTGACAGCTTCTCAAAAGAAATGGGTACAAACCCAGACACAT 1257
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1258 AAGGAAATTTGGAATTCAAAATATTCACTTCAGTTACCCATCTCGAAAGACGTTTCAG 1317
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1318 ATCTTGAGGGCTCACTGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1378 AGTGGCTGTGGAAAGACCACTGTCCAGCTGTGCAGAGGCTCTACGACCCACAGAG 1437
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1438 GCGGTGGTTCAGTATCGAGCGTCAAGACATCAGGACCAATCAATGAGGTATCTCGGGAA 1497
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1498 ATCATTTGGGTCTGAGTCAGAACCTGTGTGTGTGCCACCACTATCGCTGAAACATT 1557
Qy 284 LysTyrGlyArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1558 CGCTATGGCCGAGAAAATGTCACTATGACGAGATCGAGAAAGCTGTCAAGGAAGCCAA 1617
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1618 GCCATGACTTCATCATGAACCTGCCCAATTTGACACTCTGTGTGGTGGAGAGGA 1677
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 1678 GCACAGCTGAGTGGAGGACAGAAACAGAGATCGCCATTGCTCGTCCCTGGTCCGCAAC 1737
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 1738 CCCAAGATCTTTTGTGATGAGCGACATCAGCTTTGGACACAGAAAGTGAAGCCGCTG 1797
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 1798 GTTCAGGCTGCTCTGGTAAAGCTAGAGAGCGCGGACTACCATTTGTGTAGTCCACCGC 1857
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 1858 TTGCTACAGTTTCGAATCTGTACATCATAGCTGGTTGTGATGGTGTGCTCATTTGGAG 1917
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 423 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
```



Db 655 AAGCGGATCAAGCCAAACATTTCCATGGGCGCAGCTTCTCTTATCTATGATCATAT 714  
QY 104 GlyLeuAlaPheTptTyGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyThrIle 123  
Db 715 GCCTGGACCTTCGTATGGAACGTCCTTGGTCATCTCAAAAGBA-----TATCTATT 768  
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyCysIleGlyAlaAla 143  
Db 769 GGACAGTGCTCACTGCTCTCTTGTGTATTAATTCGGCCATTCAGTATGGACAGCA 828  
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 829 TCTCAATATCGAAGCCTTTCGCAATGCAAGAGCAGCCTATGAAATCTTCAACATA 888  
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyIleProGluSerIle 183  
Db 889 ATTGATTAATAGCCAGATTGACAGCTTCTCAAGAAATGGGTACAAACCCAGACAACT 948  
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyProSerArgProSerIleLys 203  
Db 949 AAGGAAATTTGGAATTCAAAATATTCACCTTCAGTTACCCATCTCGAAAAGACGTTGAG 1008  
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1009 ATCTGAAGGGCCTCAACCTGAAGGTGACAGCGGACAGACAGTGGCCCTGGTTGGCAAC 1068  
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAsp 243  
Db 1069 AGTGGCTGTGGGAAAGACCACTGCTCAGCTGCTGACAGAGCTCTACGACCCACAGAG 1128  
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp 263  
Db 1129 GCGGTGCTAGTATCGACGACAGGACATCAGGACCATCAATGAGGTATCTGCGGGAA 1188  
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1189 ATCAATGGGGTCGTGAGTCAGGAACCTGTGTGTTCGCCCACTATCGCTGAAACATT 1248  
QY 284 LysTyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1249 CGCTATGGCCGAGAAATGTCACTATGACGAGATCAGAAAGCTGTCAGGAAGCCAT 1308  
QY 304 AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323  
Db 1309 GCCTATGACTTCATCATGAAACTGCCCCATAAATTCACACTCTCTGTGTGAGAGAGGA 1368  
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1369 GCACAGCTGATGGAGACAGAAACAGAGAAATCGCCATTCGTCGCCCTGGTCCGCAAC 1428  
QY 344 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1429 CCCAAGATCCCTTTGTTGGATGAGCGACATCAGCCTTGGACAGAAAGTGAAGCCGTG 1488  
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 383  
Db 1489 GTTCAGGCTGCTCTGGATAAGCTAGAGAGCGCGACTACCATTTGTATGATGCCACCGC 1548  
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1549 TTGTCTACAGTTCGAAATGCTGACATCATAGCTGGGTTCGATGGTGTGTCATTTGGAG 1608  
QY 404 LysGlyAlaHisAlaGluMetAlaLysArgGlyLeuTyTySerLeuValMetSer 423  
Db 1609 CAAGGAAATCATGAAGACTCATGAGAGAGAGGAGGCAATTTACTTCAAACTTGTCATGACA 1668  
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMet 433  
Db 1669 CAGACAGCAGGAAATGAAATTTAGAAATAGGAATGGAAGTGTGTGAGTCTAAAAATGAAATT 1728  
QY 434 GluSerMetThrTySerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451  
Db 1729 GATAAATTTAGACATGCTCTCAAAAGATTTCAGCATCCAGTCTTAATTCAGAGGAGATCAACT 1788

QY 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470  
Db 1789 CGCAGAAGTATCGTGGACCATGACCAAGACAGGAAGCTTAGTACCAAGAGCGCTTG 1848  
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 490  
Db 1849 GATGAAGATGTATCACTCCCAATTTCTTTTGGCGTATCTTAAAGTTGAATTCATCTGAATGG 1908  
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 1909 CCATATTTTGGTGGTATATTTCTGTCATAGTAATAATGGAGCCTTGCACAGCATTC 1968  
QY 511 SerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysTyThrLeu 529  
Db 1969 TCAATAATATTTCTAAGGTGTGAGGGTTTTCACAAGAAATACTGATGATGAACCAAA 2028  
QY 530 LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2029 CGACATGATAGCAACTGTTTTCTTTATTTCTTCTGATCTTGGGTCTATTTCTTTAT 2088  
QY 550 SerTyPheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2089 ACATTTTCTTCAGGCTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGACTC 2148  
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyArgGlnAspIleAlaTyPheAspGluLysGlu 589  
Db 2149 CGATACATGTTTTCATCCATCTGCTGACACAGGATGTGAGCTGGTTTGATAACCCCTAA 2208  
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2209 AACACACTGGAGCATTTGACCACAGCTCGCCACAGATGCTGGTCAAGTTAAAGGGCT 2268  
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2269 ACAGGACCAAGCTCTGTCATTTACCAATATAGCAATCTTGGAGAGAAATCATC 2328  
QY 630 IleSerPheIleTyArgGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2329 ATATCCCTAATCTATGCTGGCAGTTGACACTTCTACTCTTAGCAATTTGCCCATCAT 2388  
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2389 GCATAGCAGGAGTGTGTGAGATGAAATTTGTTCTGGACAAGACTAAAGATAAGAAG 2448  
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2449 GAGCTAGAAGTTCTGGGAGATTTGCTAGAGCAATAGAGAACTTCGCACTGTGCTC 2508  
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyArgGluMetLeuGlnThrGlnHis 709  
Db 2509 TCTTTGACTCGGAGCAGAGATTGAAATATGTATGCCAGAGCTTGCAGATACCATAC 2568  
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyArgPheSerHisAlaPhe 729  
Db 2569 AGAAATGCTCTGAAGAAAGCACACGCTTTTGGGATCACATTTCTCTCCACCCAGGCCATG 2628  
QY 730 IleTyPheAlaTyAlaAlaGlyPheArgPheGlyAlaTyIleIleGlnAlaGlyArg 749  
Db 2629 ATGATATTTTGGTATGCTGCTTTCGCGTTTGGTGTCTTACTTGTGTGGCAGCGGAATTA 2688  
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyArgGlyAlaMetAlaIle 769  
Db 2689 AUGACATTTGAAATGTTCTATTAGTATTTCTCAGCTATTGCTTGTGGTCCATGGCAGTT 2748  
QY 770 GlyLysThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaHisLeu 789  
Db 2749 GGGCAGGTGAGTTCAITTTGCTCTGACTATGCCAAAGCCAAAGTGTGCGCATCGCATATC 2808  
QY 790 PheAlaLeuLeuGluLysProAsnIleAspSerArgSerGlnGluGlyLysPro 809  
Db 2809 ATCATGATCATTTGAAAGTTCCTTCTCATTTGACGCTTACAGCACCGGAGCGCTGAAGCCT 2868

Qy	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829
Db	2869	AATAACATTCGAGGAAATGAAATTTAATGAAGTCGTGTTCACATACCCACCCGCGCT	2928
Qy	830	AspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrValAlaPhe	849
Db	2929	GACATCCAGTCGTCAGGGCTGAACCTGGAGGTGAAGAGGCCAGACGCTGGCCCTG	2988
Qy	850	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
Db	2989	GTGGCAGCAGTCGTCGGGAAGACACACAGTCGTCCAGCTGCTCGAGCGGTTCTACGAC	3048
Qy	870	ProValGlnGlyGlnValLeuPheAspGlyValAlaAlaLysGluLeuValGlnTyr	889
Db	3049	CCCATCGCCGAACAGGTGTTCTAGATGGCAAGAGTAACCAATAAATGTCACGTGG	3108
Qy	890	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	909
Db	3109	CTCCAGACACCTAGGCATGTGTCCAGAGCCCATCTCTGTTGACTGCACATCGCT	3168
Qy	910	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	929
Db	3169	GAGAACATCGCTATGGAGACACACCGCGTGTGTCCAGCATGAGATCGAGAGCGG	3228
Qy	930	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	949
Db	3229	GCGAAGGAGACCAACATCCACCATGTCATTCAGTCACTGCTGACAAATACACACAGA	3288
Qy	950	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	969
Db	3289	GTAGGAGACAAAGGACTCAGCTGTGCGGTGGGAGAAACACGATCGCCATGACACGC	3348
Qy	970	AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	989
Db	3349	GCCTCGTCAGACAGCTTCACATTTGCTCTGATGAGCAACATCAGCCCTGGATACA	3408
Qy	990	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1009
Db	3409	GAGAGTGAAAGGTTGTCCAGGAGCTCTGGCAAGCCAGAGAGCCGACCTGTCAT	3468
Qy	1010	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn	1029
Db	3469	GTATCGCTCACCCTGTCACCATCCAGACGACACTTGATCGGTGATTCAGAAAT	3528
Qy	1030	GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe	1049
Db	3529	GCGAAGTCAAGAGCATGCGACCCACAGCAGCTGCTGSCACAGAAAGCATCTATTC	3588
Qy	1050	LysLeuValAsnAlaGlnSer	1056
Db	3589	TCCATGTCAGTGTGACGGCT	3609
RESULT	9	CRUPGP1165	
LOCUS		Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.	
DEFINITION		Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.	
ACCESSION		M59253	
VERSION		M59253.1	GI:191154
KEYWORDS		multidrug resistance glycoprotein; p-glycoprotein.	
SOURCE		Chinese hamster, cDNA to mRNA, clone ADX165.	
ORGANISM		Cricetulus sp.	
REFERENCE		1 (bases 1 to 4304)	
AUTHORS		Devine, S.E., Hussain, A., Davide, J.P. and Melera, P.W.	
TITLE		Full length and alternatively spliced pgpi transcripts in multidrug-resistant Chinese hamster lung cells	
JOURNAL		J. Biol. Chem. 266 (7), 4545-4555 (1991)	
MEDLINE		91154265	
PUBMED		1671863	
FEATURES		Location/Qualifiers	
source		1. .4304	

/organism="Cricetulus sp."			
/db_xref="taxon:10031"			
/cell_type="fibroblast"			
/tissue_type="lung"			
gene	1. .4304		
	/gene="pgp-1"		
CDS	111. .3941		
	/gene="pgp-1"		
	/codon_start=1		
	/product="p-glycoprotein"		
	/protein_id="AAA37004.1"		
	/db_xref="GI:191155"		
	/translation="MEFEEDSGRKDKNLFKMGKSKKKEKKKDVVSFTMFRVAGM LDRLVMTGLAAIHGVALPLMLVFGDMTDSFASVGNIPTNATNATQVNASDIFG KLEEMTTAYTIGIGAGVLIVAYIQVSWFCMLAARGQIHKIRKQFTHAMQIEQFI DVIDGELNLTLDVDSKINGIGDKIMGFQAMATFFGGFTIGTFGKMLTFLVLA SPVLGSAGIWKILSSFTDKLOQAKAGAAVEVLAARTIVIAFGQKKLELYNN NLEAKRIGIKKAITANTSMGAFLLIYASVALAFWYGTSLVSKYSIGVLTVFFA VLIAPPSIGQASPNIEAFANARGAAEYFNIDNKPSIDSFSKNGYKPDNIKNLEPK NIHFSYPSRKDVQLKUNLQVSGQTVLGVNSCGGKSTTVQLLQRLDYDPTGCVVSI DQDITINRYLREIIGVSOEPLVFAITIAENIRYGRNVMTDEIKAVKEANAYD FMKLPKFDLIVGERGALSGQKQRIARALVRNPKILLDLATSDTESAEV QAALDKAREGRTTIVIAHRLSTVRNADIAGDGGIVBQGNHEELMREKGIYFKLVM TOTAGNETELGNEVESKNEIDNLMSSKDSASSLI RRRSTPSSIRGPHDQDRLSTK EALDEDPVPSFWRILKLNSEWPYVVGICAIYNGALQAPAFSIIFSKVGVFTTNT DDETRHDSNLSFLFLILGVISFTFFLQGTGKAGEILLTKRUYVFKSMLRQDVL SWFNPKNTGALTTRLANDAGVKGTGATCARLAVITQNIANLGTGIIISLIYGWOLTL LILATVPIITAGVVMKLSQALKDKLEGGSKIATEATENFTVTVSLTREQKE NNYAQSLQIPYRNALKKAHVGITFTSFQAMMYFSYACFRGAGLVARELMTFNVL LVFSALVFGMAVGOVSSFPADYAKAKVSASHIMIIEKVPISDSYSTGGLKPNTEG NVKNEVFNPTFRDIPVLOGLNLEVKKGQTLALVSGSGCKSGTQVQLLEFYDPM GTVFDGKEVNQNLQWMLRAHLIGVSPFILDCSAENIAENIAGDMSRVVSQDEIBRA KEANTHQFIESLPKYNTRVGDKGTQLSGGQKQRIARALVRPHILLDLATSDALD TESERKVOQEALDKAREGRTTIVIAHRLSTQNIADLIIVIONGKVKHEGHTHQLAQK IVFMSVVOAGAKR"		
BASE COUNT	1275 a 895 c 1031 g 1103 t		
ORIGIN	Map position 1926.		
Alignment Scores:			
Pred. No.:	3.04e-196	Length:	4304
Score:	3136.50	Matches:	598
Percent Similarity:	75.82%	Conservative:	211
Best Local Similarity:	56.04%	Mismatches:	215
Query Match:	58.85%	Indels:	43
DB:	10	Gaps:	6
US-09-873-409-4 (1-1058) x CRUPGP1165 (1-4304)			
Qy	4	SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal	23
Db	813	TCATTTACTGATAAAGAACTTCAGGCATATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC	872
Qy	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnLysGluLeuGlnArgSer	43
Db	873	TTAGCAGCCATCAGAACTGTGATTCATTGGAGGACAAAGAAAGAACTTGAA-----	926
Qy	44	PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyrLeuLeuSerCys	63
Db	926	-----	926
Qy	64	ValLeu***PheValArgTyrThrGlnAsnLeuLysAspPheGlyIleLys	83
Db	927	-----AGGTACAAACAATTTAGAAAGAGCCAAAGAGCTTGGGATAAAG	971
Qy	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	103
Db	972	AAAGCATCACAGCCAAACATTTCCATGGCGGCGAGCTTCTCTTCTATCATGATCATAT	1031
Qy	104	GlyLeuAlaPheTyrTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle	123
Db	1032	GCTCTGGCGCTTCGTGATGGAACGCTCTTGGTCATCTCAAAAGAA-----TATCTATT	1085
Qy	124	GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla	143

Db	1086	GGACAAGTGTCTACTGCTCTCTTTGCTGTATTAAATTGGCCATTCAGTATTGGACAGCA	1145
Qy	144	ValProHisPheGluThrPheAlaIleAlaAAGGlyAlaAlaPheHisIlePheGlnVal	163
Db	1146	TCGCCAAATATCAGAGCCCTTTGCCAATGCAAGAGAGAGAGCTTAAATCTTCAACATA	1205
Qy	164	IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyLysProGluSerIle	183
Db	1206	ATTGATAATAAGCCAGTATTGACAGCTTCTCAAAGATGGGTACNAACGACACAACTT	1265
Qy	184	GluGlyThrValGluPheLysAenValSerPheAenTyProSerArgProSerIleLys	203
Db	1266	AAAGAAATTTGGAAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAAAGACGTTTCAG	1325
Qy	204	IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	1326	ATCTTGAAGGCCCTCAACCTGAAGTGCAGACGCGACAGACAGATGGCCCTGGTTGGCAAC	1385
Qy	224	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAsp	243
Db	1386	AGTGGCTGTGGGAAGACCACTGTCCAGCTGCTCGAGAGGCTCTACGACCCACACAGAG	1445
Qy	244	GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyArgAsp	263
Db	1446	GGCGTGTCTAGTATCGACGACAGGACATCAGGACCATCAATGTGAGGTATCTCGCGGAA	1505
Qy	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	283
Db	1506	ATCAATTGGGGTCTGTAGTCAGGAACCTGTGTGTTGGCCACCCTATGCTGAAACCACTT	1565
Qy	284	LysTyArgLysArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
Db	1566	CGCTATGGCCGAGAAATGTCATCTAGGACGAGATCGAAGAAGCTGTCAAGAGAAGCCAT	1625
Qy	304	AlaTyArgPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly	323
Db	1626	GCCTATCACTTCATGAACTGCCCCATAAATTGACATCTTGTGTTGGTGAGAGAGCA	1685
Qy	324	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
Db	1686	GCACAGCTAGTGGAGGACAGAAACAGAGAAATCGCCATTCGTCGCGCCTGGTCCGCAAC	1745
Qy	344	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1746	CCCAAGATCCCTTTTGTGGATGAGGCGCATCAGCCTTGGACACAGAAAGTGAGCCGTG	1805
Qy	364	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383
Db	1806	GTTCCAGCGTCTCTGGATAAGGCTAGAGAAGCCGCGACTACCATTTGTATGATGCCACCGC	1865
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1866	TTGTCTACAGTTCCGAATGCTGACATAGCTGGGGTTGATGGTGGTGTGTCATCTGTGGAG	1925
Qy	404	LysGlyValHisAlaGlnLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSer	423
Db	1926	CAAGGAAATCATGAGAGCTCATGAGAGAGAGGCAATTACTTCAAACTTGTCTATGACA	1985
Qy	424	Gln-----AspIleLysLysAlaAspGluGlnMet	433
Db	1986	CAGACAGCAAGAAATGAAATTGAAATTAGGAATTAAGTTGGTGAGTCTTAAAAAATGAAATT	2045
Qy	434	GluSerMetThrTySerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer	451
Db	2046	GATAAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTTAATCAGAAAGAGATCAACT	2105
Qy	452	ValLysSerIleLysSerAspPheIleAspLysAlaGluLysThr---GlnSerLys	470
Db	2106	CGCAGAAGTATCCGTGGACACATGCCAAGACAGGAGACTTAGTACCAAGAGAGCCCTTG	2165
Qy	471	GluIleSerLeuProGluValSerLeuLeuLysLysIleLeuLysLeuAenLysProGluTrp	490

Db	2166	GATGAAGATGACTCCCAATTTCTCTTTGGCGTATCTCTAAAGTTGAATTCATCTGAATGG	2225
Qy	491	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	510
Db	2226	CCATATTTTGGTGGTATATCTGTGCCATAGTAATGGAGCCTTGCACACCGCATTC	2285
Qy	511	SerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeu	529
Db	2286	TCATAATAATTTTCTAAGGTGTAGGGGTCTTTCACAGAAATACTGATGATGAACACAAA	2345
Qy	530	LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal	549
Db	2346	CGACATCATGACCAACTGTGTTCCTTATTTCTTCATGCTCTGGGGTCAITTTCTTTATT	2405
Qy	550	SerTyrrPheMetGlnGlyLeuPheTyrrGlyArgAlaGlyGluIleLeuThrMetArgLeu	569
Db	2406	ACATTTTTTCTTCAGGCTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGACTC	2465
Qy	570	ArgHisLeuAlaPheLysAlaMetLeuTyrrGlnAspIleAlaTrrPheAspGluLysGlu	589
Db	2466	CGATACATGTTTTTCAATCCATGCTGCACACAGGATGTGAGCTGTTTGATAACCCATAA	2525
Qy	590	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	609
Db	2526	AACACCATCGAGCATGTGACACACAGGCTGCGCACAGATGCTGGTCAAGTTAAAGGGCT	2585
Qy	610	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	629
Db	2586	ACAGAGCGAGACTTGTCTCATACCCAGANATATAGCAAATCTTGGGACAGGAATCATC	2645
Qy	630	IleSerPheIleTyrrGlyTrrPgluMetThrPheLeuIleLeuSerIleAlaProValLeu	649
Db	2646	ATAATCCTAATCTATGCTGGCAGTTTGAACATCTTACTCTTAGCAATTTGTCCTCCATCAT	2705
Qy	650	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	669
Db	2706	GCAATACGACGAGTGGTGTAGATGAAAAATGTTGTCCTGGCAAGCACTAAAAAGATAAGAG	2765
Qy	670	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
Db	2766	GAGCTAGAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCGGCACTGTCTGTC	2825
Qy	690	SerLeuThrArgGluLysAlaPheGluGlnMetTyrrGluGluMetLeuGlnThrGlnHis	709
Db	2826	TCTTTTGACTCGGAGCAGAAAGTTGAAATAATGATGATGCCAGAGCTTGCAGATACCATAC	2885
Qy	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrrAlaPheSerHisAlaPhe	729
Db	2886	AGAAATGCTCTGAAGAAAGCACACGCTCTTTGGGATCACATCTCTCTCACCAGGCCCCATG	2945
Qy	730	IleTyrrPheAlaTyrrAlaAlaGlyPheArgPheGlyAlaTyrrLeuIleGlnAlaGlyArg	749
Db	2946	ATGATATTTTTCGTATGCTGCTGTTTCCGGTTTGGTGCTTACTTTGGTGGCAGCGAAATTA	3005
Qy	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrrGlyAlaMetAlaIle	769
Db	3006	ATGACATTTTGAATAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGCGAGTT	3065
Qy	770	GlyLysThrLeuValLeuAlaProGluTyrrSerLysAlaLysSerGlyAlaAlaHisLeu	789
Db	3066	GGGCAAGTTCAGTTCAATTTGCTCTCTGACTGCCAAAGCCAAAGTGTGCGCATTCGCATATC	3125
Qy	790	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	809
Db	3126	ATCATGATCATTTGAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCTTGAGGCT	3185
Qy	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrrProCysArgPro	829
Db	3186	AATACATTTGGAAGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCACCCCGGCT	3245
Qy	830	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	849
Db	3246	GACATCCAGTGTCTCAGGCGCTGAACCTGGAGGTGAAAGGGCCAGACACTTGGCCCTG	3305





QY 101 GlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120  
DB 1054 GCATCATATGCACTGGCCCTCTGGTATGGAACCTACCTTGATCTTAGCCCAATGAG----- 1107  
QY 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIle 140  
DB 1108 TACAGTATTGGCAATGTTCTTACAGTCTTTTCTGTACTGATTGGAGCTTTTCAGTATT 1167  
QY 141 GlyAlaAlaValProHisPheGlyThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
DB 1168 GGACAGACTGCTCCAAGCATAGAGCATTTGCTTAATGCAAGAGAGCAGCCTATGCAATC 1227  
QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180  
DB 1228 TTTAATATCATAGCAATGAACCTGAAATTTGACAGTTATTTCAGATGAGGCCCAAAACCA 1287  
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200  
DB 1288 GACCACATTAAAGGAATTGGAAATTCCTCAAAATGTTTCTTCAATTAACCATCCAGACCA 1347  
QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
DB 1348 GATGTTGAGATCTGAAAGCCTTGAATCTCAAGGTTAAATGGCGCAGACAGTAGCCCTG 1407  
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240  
DB 1408 GTTGGTGCGAGCGCTGTGGGAAAGTACAACTGTTTCAGCTCATCCAGAGATTTACGAC 1467  
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260  
DB 1468 CCCAAGGAAGGCAGCATTTACCATTTGCGGCGAGGATCTGAAGAGCCTGAATGTAAGATAT 1527  
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280  
DB 1528 CTGCGAGAGATTATTGGTGTGGTGAACAGGAGCGCGTGTGTTGCTACAACTATTGCA 1587  
QY 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300  
DB 1588 GAAATATTGTTATGGCCGTGAGGATGTCCATCATGGAAGAAATGGAAGAGCTACCAAG 1647  
QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
DB 1648 GAAGCTAATGCTTATGATTTTCATCATGAAGCTCCCAAGAAGTTGAAACTGTGGTTGGA 1707  
QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340  
DB 1708 GAAAGAGGGGCACAGATGAGTGGAGGGCCAGAGCAGCGAATAGCAATTGCGCGTCTG 1767  
QY 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
DB 1768 GTTCACAATCTCTAAATTTCTGCTTGATGAGGCAACGTCAGCTTTGGATACGTAGAGT 1827  
QY 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380  
DB 1828 GAATCCCTGTGCGAGGAGCAGCTGGACAAGCAGAGAGAGGAGCAGCCCGTTGTAGTA 1887  
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
DB 1888 GCTCATGCTCTATCAACAGTCCGAAATGCGATCTTATCGCTGTGTTGAAAGTGGAGTT 1947  
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420  
DB 1948 ATCACAAACAAAGGAATTCATCTCAACTGATGAAAAAAGGAATCTACTATAAACTT 2007  
QY 421 ValMetSerGlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
DB 2008 GTTAAACATGCAGACATAGAAATGGAATCCATCATCAGAAAAAGTGAATGCGATC 2067  
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe--- 459  
DB 2068 AGTGTCAAAAGAGTGGATCTCAATCAATCTGGATGAATCCCTTAAAAAAGAAATTGAGA 2127  
QY 460 -----IleAspLysAlaGluGluSer 466

DB 2128 AGAGGATCAACTAGGAGAGCATGAAAAAGCCAGAGAACCTTAATGATACCTGATGAGAA 2187  
QY 467 ThrGlnSerLysGlyLeuIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 486  
DB 2188 GGTAGTTACCTGATGAGAAATACCTCCAGTTTCATTTCTGAAACTATGAGTTGAAC 2247  
QY 487 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506  
DB 2248 AAAACGAATGCGCATACCTTTGTAGCTGGAACCTTTGTGCAATTTGTCAATGGAGCTTGA 2307  
QY 507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 526  
DB 2308 CAACCTGCATTTTTCAGTCATATTTCTCAGAAATATATAGGATTTTTCAGAAAATGATCAG 2367  
QY 527 ThrThrLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 546  
DB 2368 AAGATTTTGAGAAAAAGAGCAACTTATATTCCTGCTGCTTTTTCAGCACTTGGGATCAT 2427  
QY 547 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 566  
DB 2428 TCCTTTTACTTTCTTGTTCAGGGTTTTCATTTGGCAAGACTGGAGAAATTTCTTACA 2487  
QY 567 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAsp 586  
DB 2488 ATGAAGCTTCGATTTCATGCAATTTAAAGCAATGCTGAGACAGGACATGGCCTGGTTGAT 2547  
QY 587 GluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 606  
DB 2548 GATCTTAAAAATAGCACTGGAGCAGCTGACAAAGACTTGTAAATGCTCCACAAAGTA 2607  
QY 607 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 626  
DB 2608 AAAGAGCAACTGGTGTGAGATTGGCATTAATTTGCCCAAAACATAGCTAACCTGGGAAC 2667  
QY 627 SerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAla 646  
DB 2668 GGAATATTATATATCTTATGTTTATGCTGGCAGCTGACCTGTTACTCTTAGCTGTGTA 2727  
QY 647 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 666  
DB 2728 CCAATCATGCTGAGCGCAGGAATGATTGAAATGAAGATGTTGGCTGGCATGCCAAAAA 2787  
QY 667 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 686  
DB 2788 GATAAATAGAGCTGGAAGCTGCAGGAAAGATTGCCACAGAGCCATAGAAAAATATTCGA 2847  
QY 687 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 706  
DB 2848 ACTGTTGCTAGTTGACTCGGAAAAAAGATTGAGCTCATGTACGAGAACATTTGCTT 2907  
QY 707 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 726  
DB 2908 GTACCATACAGAAATTTCTGTGAGAGAGGGCGCATATTTGGGTTTGTGTTTCCCTTTCA 2967  
QY 727 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 746  
DB 2968 CAAGCAATGATGTTTTTACCTATGCTGCTGTTTCGCTTTGGAGCCTATCTTGTGTT 3027  
QY 747 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 766  
DB 3028 AATGGACACATACAGTACAAACCTGTTTGTAGTGTGTTTTCAGCTGTTGTTGTTGCT 3087  
QY 767 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 786  
DB 3088 ATGCGAATAGGAAACCCAGCTCTTTTGACCAAGATATGCCAAAGCCAAAGATATCAGCA 3147  
QY 787 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 806  
DB 3148 GCCACCTTGTGTTGCTGTTTAAATAGATACCTCCATAGACAGATTACCCGAGGATGGA 3207  
QY 807 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 826  
DB 807 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 826



Db	3208	GAGAAACCTGAAAAATTTGGGGAAACACACAGGATCAAGATGTGAATTTAACTATCCA	3267
Qy	827	CysArgProaspValPheIleLeuArgGlyLeuSerIleuSerIleuGluArgGlyLeuThr	846
Db	3268	AACCGACGAGGCTCAAAATCCTCCAGGCTTGAATCTAGCAGTAGAAAAAGAGAAACG	3327
Qy	847	ValAlaPheValGlySerSerGlyCysGlySerThrSerValGlnLeuLeuGlnArg	866
Db	3328	TTGGCCCTTGTGTGATGAGATGATGGAAGAGACACTGCTTCTGAGCTACTTGAGAGA	3387
Qy	867	LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn	886
Db	3388	TTTTATGATCCACTCAGTGGAGAAATGTGTTGATGACATGATGCAAGACACTAAAT	3447
Qy	887	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	906
Db	3448	ATCCAGTGGCTGAGATCTCATCGGTATCGTCTCAAGAGCCAACTCTGTTTCGACTTC	3507
Qy	907	SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle	926
Db	3508	ACCATCGCTGAAACATCGCGTACGGGACACAGTCGGGAGGTGCCCATGAGGAAATC	3567
Qy	927	LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr	946
Db	3568	ATTAGTGCAGCGAAAGCAGGACGATTTATCTCTTCACTGACTCTCTGCCGAGAAATAC	3627
Qy	947	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	966
Db	3628	AATACTCGTGTAGGACAGCAAGGCAACACAGCTTCTTGGTGTGTCAGAAACAACGTATTGCT	3687
Qy	967	IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla	986
Db	3688	ATTGCTCGAGCTTTATACGTAAAGCCCAAGATTCGTCTACTGGATGAAGCTACATCTGCC	3747
Qy	987	LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	1006
Db	3748	TTGGACAGCAAGGCAAGAAAGATTGTCAGAGACGCTGATTAAGCCGAGAGGTGCG	3807
Qy	1007	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal	1026
Db	3808	ACCTGCATCGTGATGCTCACCGCTGTCTCCACCATCCAGATGCTGACAGATGTGCTGTG	3867
Qy	1027	LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp	1046
Db	3868	ATCCAGCAAGCGCAAGGTCATAGACCAAGGACCTCACCAGCAGCTCTGCTGCTGAAAGGC	3927
Qy	1047	IleTyrPheLysLeuValAsnAlaGlnSer	1056
Db	3928	TTCTACTCTGCTGGTCAATGTTCAAAGT	3957
RESULT 11			
LOCUS	OUA78609	3858 bp	mRNA linear
DEFINITION	Ovis aries multidrug resistance protein-1 (mdrl) mRNA, complete cds.		
ACCESSION	U78609		
VERSION	U78609.1	GI:2149086	
KEYWORDS			
SOURCE	Ovis aries.		
ORGANISM	Ovis aries		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.		
AUTHORS	Longley, M., Phua, S.H., van Stijn, T.C. and Crawford, A.M.		
TITLE	Isolation and mapping of the first ruminant multidrug resistance gene		
JOURNAL	Anim. Genet. 30 (3), 207-210 (1999)		
MEDLINE	99371931		
PUBMED	10442984		
REFERENCE	2 (bases 1 to 3858)		
AUTHORS	Longley, M. and Crawford, A.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-NOV-1996) Biochemistry Department, University of Otago, Dunedin, New Zealand		
FEATURES	Location/Qualifiers		
source	1..3858		
gene	/organism="Ovis aries"		
CDS	/db_xref="taxon:9940"		
	1..3858		
	/gene="mdrl"		
	1..3858		
	/gene="mdrl"		
	/codon_start=1		
	/product="multidrug resistance protein-1"		
	/protein_id="AAB58489.1"		
	/db_xref="GI:2149087"		
	/translation="MDLEGDRNGRAAGNFKRDKRFFSKDEKKEKRPVSTFTMF		
	IRYSNWDLRLCMLVGLTAAI IHGAGLPLMWLVGDMTDFAGAGNLGNITLSINSTST		
	RDTSYKKLEKEMTTAAIYYSIGAGVLIAYIOVSFWCLAAAGQVHRIKQFFHAI		
	MQEELGWFVDVGVKLNRLNDUSKINEGIGDKIMFQAMATFTFTGFTGWN		
	LTVLIALIPVLGSLAAIWAKILLSFTDKLLAYAKAGAAVEVLAALKTVAFGGQK		
	KELERYNKLBEAKRIGIKKAITANISMAAPFLLIYASYALAFWYGTSLVLSREYSIG		
	IKGNLEFRNVHFHYPSRNEVKILKGLNLKVGSGQTVLVGSGCGKSTTVQLMQRLYD		
	QVLTVFFSVLIGAFSIGQASPNIEAFANARGAEVFKIIDNKPSIDSYSTNGHKPDN		
	PTGEMVSDIGODIRTIINVRYLREIIGVVSQEPVLPAITTAENIRYGRDVTWDELOKA		
	VKEANAYDFIMKLPNKFTLVGERAQLSGGOKORIAARALVRNPKILLDEATSAI		
	GTESAVVQAALDKARKGRTIIVIAHRLSTVRNADIIAGLDGCVIVEEGSHDELKMKR		
	GLYFKLVTMTKGNLELENTPGESKTDLLYMSQDSRSRLIRKSTRRSIRGSQS		
	QDKRLSTEETLDESVPVSWRILKNITTEWPFVVGVCALINGALOPAFSVIFSR		
	IGIFPRNDNETKRONSLFLILIGITISFTFLOQTFPKAGIELTRRLRLVLF		
	RSMLRDYVWFDDPKNTTGAITRLANDAAQVKGAVGSLAVITQNIANLGTGIISL		
	LYGWLITLLLAIVPIIIVAGVIEMLSGQALKDKKELEGAGKATATEAENFRVVS		
	LTREERFYTAQSLQVPRNSLRRAHVFGITFSIQAMWYFSYAGCFRFLGAYLVAQG		
	IMEFDVILLVFSVAVFGAMVQVSTFAPDYAKAKVSAAHVINIIEKIPIDYSYTEG		
	LKPSYVEGSAFNDVFNYPTRPDVPLRGLSLVKGQTLALVSSGCGKTVQQLL		
	ERFYDLACTVFDGKQNLNVMRAHMGIVSOEPIILFDCSICENIAYGDNVSVS		
	QEEIHAKEANIHSFIEMLPKYKTVGDKGTQLSGGOKORIAARALVRPHILL		
	DEATSAIDTESKVVQEAALDKARKGRTICIVIAHRLSTIQNADLIIVVFQNGRIKEGHTH		
	QQLAAQKGIYFTMVSVQAGTKRQ"		
BASE COUNT	1072 a	853 c	1002 g
ORIGIN			931 t
Alignment Scores:			
Pred. No.:	4,01e-195	Length:	3858
Score:	3118.50	Matches:	602
Percent Similarity:	75.30%	Conservative:	203
Best Local Similarity:	56.31%	Mismatches:	217
Query Match:	58.51%	Indels:	47
DB:	4	Gaps:	6
US-09-873-409-4 (1-1058) x OUA78609 (1-3858)			
Qy	4	SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluGluVal	23
Db	727	TCATTTACTGATAAAAACTTTTGGCATATGCAAAAGCTGGAGCAGTAGCTGAAGAGTTC	786
Qy	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
Db	787	TTAGCGGCTATTAAAACTGTGATTGCAATTTGGAGGACAAAAAAGAACTTGAA-----	840
Qy	44	PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys	63
Db	840	-----	840
Qy	64	ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	83
Db	841	-----AGGTACACAAAACTTTGGAAGAGCCAGAGAAATTTGGAATAAG	885
Qy	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	103
Db	886	AAGGCTATCAGGCCCAACATCTCTATGGAGCTGCTTTCTTGTGTATCTATGATCCTAT	945
Qy	104	GlyLeuAlaPheThrTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle	123
Db	946	GCCTGGCGTTCTGGTACGGGACCTCTCTTGGTCTCTCAAGAGAA-----TATTCCATT	999

Qy	124	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAa	143
Db	1000	GGACAGTGTCTACAGCTCTCTTTCTGATTAATATGAGCTTTTAGTATTGCACAGCA	1059
Qy	144	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	163
Db	1060	TCTCCAAACATTGAGGCATTGTCAAACGCAAGAGAGCAGCTTATGAAGTCTTCAAGATC	1119
Qy	164	IleAspIysIysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	183
Db	1120	ATTGATAACAAACCAAGCATTCGACAGCTATTCAAACACTGGGCAAAACCGCAATATT	1179
Qy	184	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
Db	1180	AAAGGAATCTGGAAATTCAGAAATGTTCACTTCCATTAACCATCTCGAACCAAGTTAAG	1239
Qy	204	IleLeuIysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	1240	ATCTTTGAAGGCTCAACCTGAAGGTAGGAGCGGCGACAGCGTGGCCCTGTTGGGAAC	1299
Qy	224	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	243
Db	1300	AGTGGCTGTGGGAAGACCAACCGTCCAGCTGATGCAGAGGCTGTATGACCCACGGAG	1359
Qy	244	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	263
Db	1360	GGCATGTGTAGTATCAGGACAGGAATATCAGACCATCAAGTAAAGTATCTCGGGAG	1419
Qy	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	283
Db	1420	ATTATCGGGTGTGAGTCAGGAGCCTGTGCTGTTGCCACCAGATAGTCAGAACATT	1479
Qy	284	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaArgGluAlaAsn	303
Db	1480	CGCTACGGCCGTCAAGATGTCCACATGGATGAGATTCAAAAGCTGTGAAGGAAGCCAAC	1539
Qy	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323
Db	1540	GCCTATACCTTTATCAAGAACTGCCCTAATAGTTTGACACCTGGTTGGAGAGAGGA	1599
Qy	324	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
Db	1600	GCCAGCTAAGCGGTGACAGAGCAAGAAATCGCCATTTGCCGGCCCTGTTCCGCAAC	1659
Qy	344	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1660	CCCAAGATCTCTCTGTTGGATGAGGCCACTCAGCGCTGGGCCACGAGAGTGAAGCGGTG	1719
Qy	364	ValGlnAlaIleGluGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383
Db	1720	GTTTCAGCGCCCTGGTATAGGCCAGAAAGCCCGGACCAACCATTTGATAGCCATCGT	1779
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1780	CTGTCCACGGTGTGCAATGCTGACATCATTTGCTTGTGATGATGGAGTCATCGTGGAG	1839
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	423
Db	1840	GAAGGCAGTCATGATGAGCTCATGGGGAAGAAAGAGGCATTTTACTTCAAACCTGTCACAATG	1899
Qy	424	GlnAspIleLysIysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys	443
Db	1900	CAG---ACAAGAGAAATGAACCTTGAAATTTGGAATAACCCCTGGTGAATCCCTAAGTAAA	1956
Qy	444	ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle-----	460
Db	1957	ACTGATGACTTGTACATGTCTCTCCAGATTCA---AGATCAGTCTTAATCAGAGAAAA	2013
Qy	461	-----AspLysAlaGluGluSerThrGln	468
Db	2014	TCAACTCGCAGGAGTATCCGTGGATCCCAAGCCAGGACAGAAAGCTCAGTACAGAGAA	2073

Qy	469	SerLysGluSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro	488
Db	2074	ACTTTGGATGAAGTGTACTCCCTGTTTGGAGAGATTCTGAACCTGAMATAACT	2133
Qy	489	GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	508
Db	2134	GAATGGCCATTATTGTGGTGGTGATATTTTCGGCCATTATAAATGGAGCTCTGCAACCA	2193
Qy	509	ValPheSerIlePheAlaLysIleThrMetPheGlyAsnAspLys---Thr	527
Db	2194	GCATTCTCAGTCATATTTTCAGGATTATAGAACTCTCCACGAGAAATGACAAACACGAA	2253
Qy	528	ThrLeuLysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCys	547
Db	2254	ACCAAAACGACAGAACCACTGTTTTCACCTGTTGTTTCAATCCTTGGAAATATTCT	2313
Qy	548	PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet	567
Db	2314	TTTATTACATTTTTCTTCAGGCTTCACGTTTCGGCAAGCAGGCGAGATCCTCACCAGG	2373
Qy	568	ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu	587
Db	2374	CGCTCGGTACTCTGTTTTCAGTCCATCTGACAGCAGATGTCAGCTGGTTCGATGAC	2433
Qy	588	LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln	607
Db	2434	CCTAAAAACACCCAGGAGCATTTGACGACACGCGTTGCCAATGACCGCTCMAGTAAA	2493
Qy	608	GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer	627
Db	2494	GGGCTGTAGGTTCCAGGCTGCTGTAATTATCCCAAGATTTGCCAAATCTTGGCACAGGC	2553
Qy	628	ValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaPro	647
Db	2554	ATTATCATTCCTCATCTATGCTGGCAGTTAACACTTCTCCTGTTGCAATTTGATACC	2613
Qy	648	ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp	667
Db	2614	ATCATTCAGTAGCAGGAGTTATTGAAATGAAATGCTGCTCTGGACAAGCCCTGAAAGAT	2673
Qy	668	LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr	687
Db	2674	AAGAAAGAACTGGAAGGTGCTGGGAAGATTGGCCCTGAAGCAATAGAAACTTCCGGACG	2733
Qy	688	IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr	707
Db	2734	GTGTGTTCTTTGACTCGGAGGAGGTTTGAATATATATGATGCCAGAGTTGCGAGTA	2793
Qy	708	GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis	727
Db	2794	CCATACAGAACTCCTTGAGGAGCGCATGCTTTTGGAAATTTACATTTCCATCATCTCAG	2853
Qy	728	AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGlnAla	747
Db	2854	GCAATGATGATTCTCTATGTGCTGTTTCCGGGTTTGGTGCTACTGTGGTGCCCAA	2913
Qy	748	GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet	767
Db	2914	GGCATATGGAGTTTCAGGATGTTCTCTTAGTGTTCTCAGCTGTGTCTTTGGTGCCATG	2973
Qy	768	AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla	787
Db	2974	GCTGGGACAGGTCAGTACATTGCTCCTGACTATGCCAAGGCCAAAGTGTGACAGCC	3033
Qy	788	HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluCylLys	807
Db	3034	CACGTCTCATATATCATTTGAAAAAATTCCTCTGATTGACAGCTACAGCAGGAAGCCCTA	3093
Qy	808	LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys	827
Db	3094	AAGCCGAGTACAGTGGAGAGAGTGGCATTTAATGACGTCTGTGTTCACTACCCACT	3153
Qy	828	ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal	847

Db	3154																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												</
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

FEATURES	source	Location/Qualifiers
gene		1. .4927
CDS		86..3904
		/organism="Rattus norvegicus"
		/strain="Wistar"
		/db_xref="taxon:10116"
		/sex="male"
		/cell_type="hepatocyte"
		/tissue_type="liver"
		/dev_stage="adult"
		1. .4927
		/gene="Pgyl"
		86..3904
		/gene="Pgyl"
		/note="ATP-binding cassette transport protein; Abcb1a"
		/codon_start=1
		/product="multidrug resistance protein 1a"
		/protein_id="AAF69007.1"
		/db_xref="GI:7739773"
		/translation="MSEEDLGNRADKNFSKMGKKSKKKEKKPAVSLVLTMRFYAGW LDRFYMLLGLTAAIHGIALPLMLVFGDMTDSFANVGNRRSMFSFNATDIYAKLDE MTTYAYTIGAGVLIVAYQSLMLAAGRQIHKIRQKFFHAIINQDEIGWDFYADLE GELNTLITDDVSGINIGDKIGMFGQAMATFGGIIGTRGKWLTLVLAISVLG LSAGIKWILSSPTDELOQAKAGAAVEELAAIRTVIAFGGOKLELRYNNLSA KRUGIKKAITANISMGAEFLIYAVSALAPWCTLSVSKYETIGQVLTVPFSLVGA FSVQASPNIETAFNARGAAAYEVSIIIDNKPSIDFSKSKHKPDNLQGNLEFKNIHFS YPRKQDVLKGLNGKGTVALVGNSGCGKSTTVQLLQRLYDPIEGEVSIDGQDI RTNRYVLRBIIIGVSOEPLVFATTAEINIRYGRNVTWDEIEKAVKEANAYDFIMKL PHKEDTILVGERGAQLVGQOKRIATARALVRNPKILLDEATSDLTSEAVVTQALD KAREGTTIVAHRLSTVRNADVIAGFDGVIYVEQGNHDELMREKGIYPKLVMTQTGA NEIELGNEACESKDGINVDVMSKSGSLIRERSTRKISIRGPHDODGELSTKEALD DVPSPFWRLLKUNSTWYFVFCALINGGLQAPFIIIPSKYGVGVTKNDTPELQ RONSNFIISLFLTGIIISFTFFLQGTGFKAGLEILTALRURYVFKMLRQDIIISWFDI PKNTGALTATRLMNDAAQVKGATSRVATQNIANLGTGIIISLIYQGTLLLLLA VPIIATAGVVMELMSQALKDKLELGGSKIAETAIENRPTVSVLTREQKFTMTAQ SLOIPYRNALKSPFTGTFSTQAMMYPFAACFRGAVLVAEELMTFNVLVLFSA IVFGMAVGVSVFADYAKVYASHIIRIIRIKPIDSVSTEGLEKPNMLEGNVKEN GVNFTPTRNIPVLOGLSLEVKKQTALVSGSGGKSTVQQLLERFYDPMAGTVFL DKEIILQNLQWLRNHLGVISQBPILFDCSIAENIAYGDNRSVVVSHEEIVKAAKEANI HQEIDSLPERYNTRVSGDKTQSGGQKRIATARALVRPHILLDEATSDLTSESK VSEALDKAREGRTCTIVIAHRLSTIQNADLIIVTVQNGVKEHGTHQOQLLAQKGIYFSM VSVQAGAKRS"
BASE COUNT	1474 a 1037 c 1190 g 1226 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	7.34e-195	Length: 4927
Score:	3116.50	Matches: 601
Percent Similarity:	75.73%	Conservative: 207
Best Local Similarity:	56.33%	Mismatches: 216
Query Match:	58.47%	Indels: 43
DB:	10	Gaps: 7
US-09-873-409-4 (1-1058) x AF257746 (1-4927)		
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23		
Db 773 TCATTACTGATAAGGAACCTCAGGCTTATGCAAAAGCTGAGCAGGTGCTGAAAGATC 832		
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgSer 43		
Db 833 TTAGCAGCCATCAGCAACTGTGATTCCTTTTCGAGGACAAAAGAGAACTTGAA----- 886		
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63		
Db 886 ----- 886		
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83		
Db 887 -----AGTCAATAACAAATTTGGAAGAGCTAAAGGCTTTGGGATAAG 931		
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103		
Db 932 AAAGTATCAGCGCAACATTCATCGATGGGTGCAGCTTTCTGCTTATCTATGCATCATAT 991		

Qy	104	GlyLeuAlaPheTrpTyrGlyThrSerLeuLeuAenGlyGluProGlyTyThrIle	123
Db	992	GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA-----TACACTATT	1045
Qy	124	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyCysIleGlyAlaAla	143
Db	1046	GGACAAGTGTCTACTGTCTTTTCTGTATTAAATGAGACATTGAGTGTGGGCGAGCA	1105
Qy	144	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	163
Db	1106	TCTCCAAATATTGAAGCCATTGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA	1165
Qy	164	IleAspLysLysProSerIleAspIlePheSerThrAlaGlyTyTyLysProGluSerIle	183
Db	1166	ATTGATAAATAAGCCCAAGTATAGACAGTCTCTCAAAAGAGTGGGCGACAAACCCGACAA	1225
Qy	184	GluGlyThrValGluPheLysAenValSerPheAenTyProSerArgProSerIleLys	203
Db	1226	CAAGGAAATTTGGAAATCAAAAATATTCACTTCAGTTACCGTCTCGAAAAGACGTTGAG	1285
Qy	204	IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	1286	ATCTTGAAGGCGCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGTTGGCAAC	1345
Qy	224	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAsp	243
Db	1346	AGTGGCTGTGGGAAAAGCAACATGTCTCAGCTGTCTGCAGAGGCTCTACGACCCCATAG	1405
Qy	244	GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyArgAsp	263
Db	1406	GGCAGGTTCAGTATCGACGGCAGGACATCAGACCATCAATGTGAGGTATCTCGGGGAA	1465
Qy	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	283
Db	1466	ATCATTGGGGTGTGAGTCAGGAAACCGGTGCTTTGCCACCACCAATTCGCCGAAACATT	1525
Qy	284	LysTyrcGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
Db	1526	CGCTATGGCGCAGAAAACGTCCACATGGATGAGATAGAGAAGCTGTCCAGGAAGCCAA	1585
Qy	304	AlaTyArgPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly	323
Db	1586	GCCTATGATTTCATATGAAACTGCCCCCAAAATTTGACACCTGTTGGTGAGAGAGGG	1645
Qy	324	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
Db	1646	GCGCAGCTGAGTGGGGACAGAAACAGAGGATCGCCATTGCCGGGGCCCTGGTCGCGAAC	1705
Qy	344	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1706	CCCAAGATCTCTTTTGTGATGAGGCCAGTCAAGCTTGGACACAGAAAGCGAGCCGTG	1765
Qy	364	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383
Db	1766	GTTCCAGCCGCTCTGGATAAGGCTAGAGAAGCCGGACCACTTGTATGATAGCTCACCGC	1825
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1826	TTGTCTACAGTTCGCGCAATGCTACGCTTCATTGTGTTGATGGTGGTGTCTGTGGAG	1885
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerSerLeuValMetSer	423
Db	1886	CAAGGAAATCATGATGAGTCTATGAGAGAGAAGGAATTTACTTCAAACACTTGTCTAGCT	1945
Qy	424	Gln-----AspIleLysLysAlaAspGluGlnMetGluSer-----	435
Db	1946	CAGACACAGGAAATGAAATTTGAATTAGGAATGAAGCTTGTGAATCTAAAGATGGAATT	2005
Qy	436	-----MetThrTyArgSerThrGluArgLysThrAenSerLeu-----ProLeuHisSer	451
Db	2006	GATAATGTGCATGTCTTCAAAAGATTCCAGGATCCAGTCTAATAAGAAAGATCAACT	2065

Qy	452	VallysSerileysSerAppheileAspIysAlaGluGluSerThr---GlnSerIys	470
Db	2066	CGCAAAGCATCGTGGGCCACATGATCAAGACGGGGAACCTAGCACCAAGAGCGCTCTG	2125
Qy	471	GluIleSerLeuProGluValSerLeuLeuIysIleLeuLysLeuAsnLysProGluTrp	490
Db	2126	GATGACGACGTACTCCACGCTCTCTTTGGCGGATCTCTGAAGTTGAATTCAACTGAATGG	2185
Qy	491	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	510
Db	2186	CCCTATTGTGGTGTGATTTTGTGCCATAATAATAGGAGCTTGCAACCAAGCATTC	2245
Qy	511	SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---	529
Db	2246	TCCATAATATTTTCAAAGTTCTAGGGGTTTTTACAATAAATGACACCCCTGAAATCCAG	2305
Qy	530	LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	549
Db	2306	CGGCAAGAACAGCAACTGTCTTTATTTGTTCTCGATCTCTTGGGATCATCTCTTTCAAT	2365
Qy	550	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	569
Db	2366	ACGTTTTCCTTCAGGCTTCACATTTGGCAAGCTGGAGAGATCCTCACCAAGCGACTC	2425
Qy	570	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu	589
Db	2426	CGATATATGTCCTTCAAATCCATGCTGACACAGGACATAAGCTGTGTTGATGACCTAAA	2485
Qy	590	AsnSerThrClyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	609
Db	2486	AACACACACAGGAGCGGTGACACACAGGCTTGCAATGACGTGCTCAAGTGAAAGGGCT	2545
Qy	610	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	629
Db	2546	ACAGGCTTAGGCTTCTGTTATTATCCACAGAACATAGCAATCTTGGGACAGGCATCATC	2605
Qy	630	IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	649
Db	2606	ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCTTAGCAATTTGTTCCCATCAT	2665
Qy	650	AlaValThrClyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	669
Db	2666	GCTATACAGGAGTGTGAAATGAAATGTTGTTCTGCAAGCGCTGAAAGATAAGAAAG	2725
Qy	670	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
Db	2726	GAACTAGAAGTTCCTGGGAAGATCGCTACAGAACCAATTGAAACTTTCGCACACTGCTGT	2785
Qy	690	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	709
Db	2786	TCITTTGACTCGGAGCAGAAGTTGAAACTATGTTATGCCAGAGCTTCAGATACCATAC	2845
Qy	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	729
Db	2846	AGAAATGCTTTGAAGAAGCGCACGCTCTTTGGGATCACCTTTCTCCTCACCCAGCGCATG	2905
Qy	730	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	749
Db	2906	ATGTATTCTTCCTATGCTGTTGTTTCCGGTTTGGTGCTACTTGTGGGCGAGAACTC	2965
Qy	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	769
Db	2966	ATGACATTTGAAATGCTCTGTATTAGTATTCTCAGCTATTGCTTTGGTGCCATGCGATG	3025
Qy	770	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	789
Db	3026	GGGCAAGTCATTCGCTCCTACTACGCGAAGCAAAAGTCTCGGCATCCCAATC	3085
Qy	790	PheAlaLeuLeuGlyLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro	809
Db	3086	ATCAGGATCATTCAGAAAAATCCCTGAGTTGACGCTACAGCACGAGGGGCTTGAGGCT	3145
Qy	810	AspThrCysGluGlyValAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829



[illegible]









Db	3521	ACCAGAGTAGGAGCAAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTTGCCATA	3580
Qy	968	AlAArgAlAlaLeuLeuGlnLysProLysLleLeuLeuLeuAspGluAlaThrSerAlaLeu	987
Db	3581	GCTCGCGCTCTGTAGACAGCCTCATATTTGCTTTTGGATGAGCTACATCAGCTCTG	3640
Qy	988	AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaAaThrGlyArgThr	1007
Db	3641	GATACACAAAGTGAAGAGGTTGTCCAAGAGCCCTGGCAAAAGCCAGAGAAGGCCGACCC	3700
Qy	1008	CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuValValLeu	1027
Db	3701	TGCATTGTGATCGCCACCCTGTCTCCACCATCCAGAAATGCAGATTTAATAGTGGTGT	3760
Qy	1028	HisAsnGlyLysLleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle	1047
Db	3761	CAGATGGCAAGTCAGAGGAGCATGGCACATCAACAGCTGTGGCCCAAGAGGCATC	3820
Qy	1048	TyrPheLysLeuValAsnAlaGlnSer	1056
Db	3821	TATTTTCCATGGTCAGTCTCCAGGCT	3847
RESULT 15			
AX105078			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
dog.			
Canis familiaris			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
1 (bases 1 to 4279)			
Stockert,P.J., Stenmel-Crespi,D.T., Crespi,C.L., Reif,T.C. and			
Patten,C.J.			
P-glycoproteins and uses thereof			
Patent: WO 0123540-A 22 05-APR-2001;			
GENTEST CORPORATION (US)			
FEATURES			
source			
CDS			
Location/Qualifiers			
1..4279			
/organism="Canis familiaris"			
/db_xref="taxon:9615"			
17..3862			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAC37721.1"			
/db_xref="GI:13921229"			
/translations="MDPEGKRGSAEKNFWKMGKSKKNEKKEKKTPTVSTFAMFPRYSN			
WLDRLVMLVCTMAAIITHGALPLMLVFGNMTDSFANAGSRNKTTPVIINESIITNTT			
QHFNHLEEMTYYAYYSIGAGLVAAIYQVFWCLAAGROILKIRQOFFHAIIMRO			
EIGFVDHVGELNTRTDVSKINEGDKIGWFFQSIATFFFTGIVTAVFGQKGLT			
VLIAISFVLGSAAIWAKILSPTKDELLAYAKAGAAVEELAAITRTVIAFGQKGL			
ERNKNLEAGKGIKKAITANISIGAAFLIIYASVALAFWYGTSLVISEYIGVL			
TFVPSVLIGAFSIGQASPTAEANARGAAVEIPKIDNKPISIDSKSGHKPDNKG			
NLRFKNVPSYPSRKEVILKGLNLKVQSGQTVALVNSGCGKSTTVQLMQLYDPT			
GMVICDGDIRITNVRLREITGVWSQEPVLPAITTAENIRYGRNTHDEIEKAYKE			
ANAYDFIMKLPNKPTDLEBGRQAQSGGQKQRIARAIARLVNPKILLLDTSALDTE			
SEAVVQVLDKARKGRTTIVIAHRLSTNRNADVPDGGIVKEKGNHDSLMKEKGI			
FKLVTQTRNGEILEENATGESDLEMSQDSSSLIKRRSTHRSIHPAQGDOR			
KLGTGDLNKPVPSPFWRLKLVNSTEWPFVVGIVFCALINGGLQPAFSIIFSRIGI			
FTDEQDETQRKNNSVFLVLKLGIIISFTFFLQGTFCGKAGELTKRLRYWVFRSM			
LRQVSWDFDKNTGALTTRLANDAAQVKAIGSRILAVITQNIANGTGIIISLIYG			
WQTLULLLALPIIATAGVVMKLSQALKKKELEGAKIATEIENFRTVVSLTR			
EQKFYEMAYASQLVRSQRKAHIFGVSESIITQAMMYFYAGCFRFGAYLVANEME			
FQDVLIAIFVGCANAVGQVSFPADYAKVSAAHVIMIEKSPILIDSYSHPGLKP			
NTLEGVNTFVNEVYVPTRPDIPLVQLGLEVKKGOTIALVSGSGCGKSTVQLLERR			
YDPLAGDVLIDGKEIKHLNVMWLRHLGIVSQPILFDCSIENIAGDNSRVVSHDE			
IMQAAXEANIHHFIETLPEKINTVRGDKVQLSGGQKQRIARALVRQPHILLDEA			
TSLADTSESEKVVQEAQDKAREGRTTIVIAHRLSTNRNADLIIVFQNGKVKHEGHTHQA			
LAQKGYFSVMVQAGAKR"			
BASE COUNT	1295 a	833 g	

ORIGIN

Alignment Scores:  
 Pred. No.: 1,13e-194 Length: 4279  
 Score: 1112.50 Matches: 601  
 Percent Similarity: 75.02% Conservative: 201  
 Best Local Similarity: 56.22% Mismatches: 220  
 Query Match: 56.22% Indels: 47  
 DB: 6 Gaps: 6

US-09-873-409-4 (1-1058) x AX105078 (1-4279)

QY	4	SeLLeuThrSerLysGluLeuSerLaTyrSerLysAlaGlyAlaValAlaGluGluVal	23
DB	734	TCATTTACTGATAAAAGAACTCTTGGCCCTATGCAAAAGCTGAGCAGTAGCTAGTGAAGAATC	793
QY	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
DB	794	TTAGCAGCAATCAGAACTGTGATTCCTTTGGAGGACAAAGAAGAAACTTGAA-----	847
QY	44	PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys	63
DB	847	-----	847
QY	64	ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	83
DB	848	-----AGGTACCACAAAAATTTAGAAAGAAGCTAAAGGAATTTGGGATAAAG	892
QY	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	103
DB	893	AAAGCTATCACGGCCCAACATTTCTATTGGTGGCGCTTCTTATTGATCTATGTCATCATAT	952
QY	104	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	123
DB	953	GCTCTGGCTTCTGGTAGGGACCTCCTTGGTCCTCTCCAGTGAA-----TATTCTATT	1006
QY	124	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	143
DB	1007	GGACAAGTACTCACTGTCVCTTTCTGTATTAAATGGGGCTTTTAGTATTGGACAGCA	1066
QY	144	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	163
DB	1067	TCCCCAAGCATTTGAAGCATTTGCAACCGCAAGAGGAGCAGCTTATGAATCTTCAAGATA	1128
QY	164	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	183
DB	1127	ATTGCAATAAACCAAGCATTCAGACTATTCCAGAGAGTGNCATAAACAGATAAATATT	1186
QY	184	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
DB	1187	AAGGGAAATTTGGAATTCAAAATGTTCACTTTCAGTTACCCTTCTCGAAAAAGAAATTAAG	1246
QY	204	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
DB	1247	ATCTTAAAGGGTCTCAACCTGAAGGTTCCAGAGTGGGACAGAGTGGCGCTGGTGGGAAC	1306
QY	224	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	243
DB	1307	AGTGGCTCGGGGAAGACACGACCGGTGACGTGATGACAGAGGCTCTATACCCCCACAGAT	1366
QY	244	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	263
DB	1367	GGCATGCTGTATTGATGCACAGGACATTAGCACCATAAATGTAAAGGCATCTTCGGGAA	1426
QY	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	283
DB	1427	ATTACTGGTGTGGTAGTCAGGAGCGTGTGTTGTTTGGCCACCCACGATAGCTGAAACAT	1486
QY	284	LysTyrClyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
DB	1487	CGCTATGGCGCGGAAAATGTACCATGGATGAGATTGAGAAGCTGTTTAAGGAAGCCCAAT	1548
QY	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323

BASE COUNT	1295 a	833 c	1008 q	1143 t
------------	--------	-------	--------	--------



Qy 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047  
 Db 3761 CAGATGGCAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCGCCAGAAAGGCATC 3820  
 Qy 1048 TyrPheLysLeuValAsnAlaGlnSer 1056  
 Db 3821 TATTTTCCATGGTCAGTGTCCAGGCT 3847

Search completed: March 31, 2003, 02:13:48  
 Job time : 14525 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 1034.02 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-4  
Perfect score: 5330  
Sequence: 1 MVISITKLSKAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODEL=frame+p2n.model -DB=cpn2.1/USPTO spool/US09873409/runat\_27032003.115416.19216/app\_query.fasta\_1.7544  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -IOFCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09873409 @CGN 1.1 4247 @runat\_27032003.115416.19216 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4985.5	93.5	3699	24	ABK83223	Human transporter
2	3115.5	58.5	4279	22	AAD03506	Dog P-glycoprotein
3	3112.5	58.4	4279	22	AAD03488	Dog P-glycoprotein
4	3112.5	58.4	4279	22	AAD03504	Dog P-glycoprotein
5	3112.5	58.4	4279	22	AAD03505	Dog P-glycoprotein
6	3111.5	58.4	4317	22	AAD03489	Dog P-glycoprotein
7	3106.5	58.3	4369	21	AAZ52047	Rat multidrug resi
8	3106.5	58.3	4425	21	AAZ52048	Rat multidrug resi
9	3100.5	58.2	3860	21	AAZ49332	Human wild-type mu
10	3100.5	58.2	3860	21	AAZ49333	Human G185V mutant
11	3100.5	58.2	3860	24	ABA94365	Human BCRP DNA rel
12	3100.5	58.2	3860	24	ABA94366	Human BCRP DNA rel
13	3100.5	58.2	4186	22	AAF86127	Cynomologous monke
14	3100.5	58.2	4195	22	AAF86128	Human P-gly
15	3100.5	58.2	4264	19	AAV66533	Mutated human P-gl
16	3100.5	58.2	4264	19	AAV66534	Human intestine ce
17	3100.5	58.2	4349	22	AAH57442	Human ATP binding
18	3100.5	58.2	4646	21	AAZ94738	Human mdrl gene.
19	3100.5	58.2	4646	24	AAD38994	Kidney cancer rela
20	3100.5	58.2	4646	24	ABL68592	Kidney cancer rela
21	3100.5	58.2	4646	24	ABL68880	Sequence of human
22	3100.5	58.2	4669	8	AAW70752	Hybrid vector pSF-
23	3100.5	58.2	6505	17	AAT13394	Retroviral M4 mdr-
24	3100.5	58.2	8630	21	AAZ24041	Retroviral vector
25	3100.5	58.2	8630	21	AAZ24042	Sequence of human
26	3097.5	58.1	4669	14	AAQ52726	Murine multidrug r
27	3090.5	58.0	4788	21	AAZ49335	Mouse BCRP DNA rel
28	3090.5	58.0	4788	24	ABA94368	Human multidrug re
29	3089.5	58.0	4646	15	AAQ72872	Rat mdrlb2 (multis
30	3086	57.9	4233	21	AAZ90198	Rat mdrlb2 multidr
31	3086	57.9	4233	21	AAZ27498	Multidrug Resist
32	3085.5	57.9	4378	11	AAQ04522	Human ATP binding
33	3085	57.9	3924	21	AAZ94742	Human MDR-3 DNA.
34	3085	57.9	3924	24	AAZ88971	Gene #2299 used to
35	3085	57.9	3924	24	ABN95801	Murine multidrug r
36	3085	57.9	4189	21	AAZ49334	Mouse BCRP DNA rel
37	3085	57.9	4189	24	ABA94367	Mouse multidrug re
38	3085	57.9	4313	14	AAQ38950	Human polynucleoti
39	3084	57.9	3840	24	ABL91687	Human MDR-1 DNA.
40	3084	57.9	3988	21	AAZ88973	Human P glycoprote
41	3073.5	57.7	4669	19	AAV32645	cDNA encoding huma
42	3073.5	57.7	4669	24	ABK52041	Rat sequence diffe
43	3044.5	57.1	4254	24	ABK63517	Rat sequence diffe
44	3043	57.1	3912	24	ABK63653	Human ATP binding
45	2515.5	47.2	4776	21	AAZ94744	

ALIGNMENTS

RESULT 1  
ABK83223  
ID ABK83223 standard; cDNA; 3699 BP.  
XX ABK83223;  
AC  
XX  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.  
XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;  
KW neurological disorder; muscle disorder; immunological disorder; cancer;  
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;  
KW cell proliferative disorder; cervical cancer; breast cancer;  
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;  
KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;  
KW Grave's disease; gastrointestinal disorder; Crohn's disease;  
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;  
 KW protozoal infection; helminthic infection; cardiovascular disorder;  
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

OS WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-NOV-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-250790P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

XX Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for

XX diagnosing, preventing, and treating disorders associated with an

XX abnormal expression or activity of TRICH, e.g. immunological, muscular

XX or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
 CC a recombinant polynucleotide comprising a promoter sequence operably  
 CC linked to the TRICH polynucleotide, a cell transformed with the  
 CC recombinant polynucleotide, a transgenic organism comprising the  
 CC recombinant polynucleotide, an isolated antibody that binds specifically  
 CC to TRICH, and screening for compounds which bind to TRICH, modulate  
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.

CC The polypeptides are useful for diagnosing, treating, and  
 CC preventing transport, neurological, muscle, immunological disorders  
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
 CC proliferative disorders such as cancers (e.g. leukemia, cervical or  
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
 CC catatonial), endocrine disorders (e.g. diabetes, Grave's disease),  
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,  
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
 CC other diseases and disorders detailed in the specification. They can also  
 CC be used in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of transporters and  
 CC ion channels. TRICH or its fragments may also be used in screening for  
 CC compounds that specifically bind to and modulate the activity of TRICH.  
 CC The polynucleotides can be used to create knock-in humanised animals or  
 CC transgenic animals to model human disease. The present sequence  
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

XX Alignment Scores:

Pred. No.: 0 Length: 3699

Score: 4985.50 Matches: 1002

Percent Similarity: 94.80% Conservative: 1

Best Local Similarity: 94.71% Mismatches: 2  
 Query Match: 93.54% Indels: 53  
 DB: 24 Gaps: 2

US-09-873-409-4 (1-1058) x ABK83223 (1-3699)

QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20  
 Db 682 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGGCA 741  
 QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40  
 Db 742 GAAGAAAGTCTTGTCAATCCGAACAGCATAGGCTTTAGGGCCCGAGAGAAAGAACTT 801  
 QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60  
 Db 802 CAA----- 804  
 QY 61 LeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80  
 Db 805 -----AGGTATACACAGAAATCTCAAAGATGCAAGGATTTT 840  
 QY 81 GlyIleArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100  
 Db 841 GGCATAAAAGGACTATAGCTTCAAAGGTCTCTTTGGTGTGTGTACTTCTTTATGAAT 900  
 QY 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120  
 Db 901 GGNACCTATGGACTTGCTTTTGGTATGGAACCTCTCTGATCTTAATGGAGAACCTGGA 960  
 QY 121 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 140  
 Db 961 TATACCATCGGACTCTTCTGCTGTTTCTTAGTGTAAATCCATAGTAGTATTGCATT 1020  
 QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
 Db 1021 GGAGCAGCAGTCCTCCTCACTTTGAACCTTCGCAATAGCCGAGGAGCTGCTTTCATAT 1080  
 QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180  
 Db 1081 TTCCAGGTATTGTAAGAAGAACCCAGTATAGGTAACTTTCCACAGCTGGATATAACCT 1140  
 QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200  
 Db 1141 GAATCCATAGAAGAACTGTGGAATTTAAATAATGTTCTTCTCAATTTATCCATCAAGGCCA 1200  
 QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
 Db 1201 TCTATCAAGATTCTGAAGGCTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCCTTG 1260  
 QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240  
 Db 1261 GTCCGCTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATGAT 1320  
 QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260  
 Db 1321 CCGGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGGCAT 1380  
 QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280  
 Db 1381 TATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCCCTGTTTGTTCGGGACCACCATCAGT 1440  
 QY 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300  
 Db 1441 AACATATCAAGTATGGACGAGATGATGTGACTGATGAGAGATGGAGAGAGAGCAAGG 1500  
 QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
 Db 1501 GAAGCAAAATGCGTATGATTTTATCATGGAGTTTCTCTAAATAAATTAATACATGGTAGGG 1560  
 QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGluArgIleAlaIleAlaArgAlaLeu 340  
 Db 1561 GAAAAAGGAGCTCAAAATGATGGAGGGCAGAAAAACAGAGGATCGCAATTGCTCTGTGCTTAA 1620



QY 341 ValArgAsnProIysLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
DB 1621 GTTCGAAACCCCAAGATTCTGATTTTATAGTAGGCTACGCTCCCTGGATTTCAGAAAGC 1680  
QY 361 LysSerAlaValGlnAlaLeuGluLysAlaSerIysGlyArgThrIleValVal 380  
DB 1681 AAGTCAGCTGTTCAAGCTGCACCTGGAGAGGCGAGCAAGGTCGAGTCAATCGTGGTA 1740  
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
DB 1741 GCACACCGACTTCTACTATTTCAGAGTCAGATTGATTGTGACCCCTAAGGATGGAATG 1800  
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeu 420  
DB 1801 CTGCGCGAGAAAGGAGCGCATGCTGAACATAATGGCAAAACGAGTCTATATTATTACATT 1860  
QY 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
DB 1861 GTGATGTCACAGGATATTAAAAAGGCTGAAGACAGATGGAGTCAATGACATATTCTACT 1920  
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460  
DB 1921 GAAAGAAAGCCAACTCACTTCTCTGCACTCTGTGAGAGCATCAAGTCAGACTTCATT 1980  
QY 461 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480  
DB 1981 GACAAGGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTCTCTGAAGTCTCTCTATTA 2040  
QY 481 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 500  
DB 2041 AAAATTTTAAAGTTAAACAAGCCCTGAATGGCTTTTGTGGTCTTGGGGAATTTGGCTTCT 2100  
QY 501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520  
DB 2101 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAATTTATAACCATG 2160  
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540  
DB 2161 TTTTGGAAATTAATGATAAAACCATTAAGCATGATGCAGAAATTTATTCATGATATTC 2220  
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560  
DB 2221 GTCATTTTGGGTGTATTTCCTTTGTTCAGTTATTTCATG----- 2259  
QY 561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580  
DB 2260 -----CAG 2262  
QY 581 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600  
DB 2263 GATATTCCCTGGTTTGTATGAAGAGAAACAGCACAGAGGCTTGACAAATATTAGCC 2322  
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620  
DB 2323 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATGGCGTCTTTAAACACAAAAAT 2382  
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640  
DB 2383 GCAACTTAACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGGGAGATGACATTC 2442  
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660  
DB 2443 CTGATTTCTAGTATTGCTCCAGTACTTTCGCTGCAGAGAAATGATTGAAACCGGCAGCATG 2502  
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyTyrIleAlaThrGlu 680  
DB 2503 ACTGGATTTGCCAACAAAGATAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2562  
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700  
DB 2563 GCTTTTGAGAAATATACGTACTATAGTGTCTATTAAACAGGGGAAAAAGCCTTCGAGCAATG 2622

RESULT 2

QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720  
DB 2623 TATGAGAGATGTTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2682  
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740  
DB 2683 AGCTGTTATGCATTTAGCCATGCTTTATATATTTTGCCTATGCGCAGGGTTTCGATTT 2742  
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760  
DB 2743 GGAGCCATTTAATTCAGCTGGACGAATGACCCACAGAGGGCATGTTTCATAGTTTACT 2802  
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780  
DB 2803 GCAATTTGCATATGGAGCTATGCCATCGGAGAAACGCTGTTTGGCTCCTGTAATATTC 2862  
QY 781 LysAlaLysSerGlyValaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAsp 800  
DB 2863 AAAGCCAAATCGGGGCTGCGCATCTGTTTGGCTTGTGGAAAAAGAAACCAATATAGAC 2922  
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820  
DB 2923 AGCCGCAAGTCAAGAGGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCAGAA 2982  
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840  
DB 2983 GTCTCTTTCTTCTATCCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 3042  
QY 841 IleGluArgGlyIysThrValaPheValGlySerSerGlyCysGlyLysSerThrSer 860  
DB 3043 ATTGACGAGGAGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCT 3102  
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880  
DB 3103 GTTCACTTCTGCAGAGACTTTATGACCCCGTCGACAGGACAGTGTGTTGATGGTG 3162  
QY 881 AspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGlu 900  
DB 3163 GATGCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCAAATAGCAATCGTTCCTCAAGAG 3222  
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920  
DB 3223 CCTGTGCTCTTCAACTGCGAGCATTTGTGAGAACATCGCTATGTCGACACACCCGTGTG 3282  
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940  
DB 3283 GTGCCATTAGATGAGATCAAGBAGCCGCAATGCAGCAATATCCATTCTTTATTGAA 3342  
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960  
DB 3343 GGTCTCCCTGAGAAATACAAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3402  
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeu 980  
DB 3403 CAGAAACAAAGACTAGCTATTGCAAGGGCTTCTTCTCAAAACCCCAAAATTTTATTGTTG 3462  
QY 981 AspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAsp 1000  
DB 3463 GATGAGCCACTTCAGCCCTCGATATGACGTGAGAGAGTGGTTTCAGCATGCCCTTGAT 3522  
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020  
DB 3523 AAAGCCAGGAGCGGAGGACATGCTGTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3582  
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040  
DB 3583 GCAGATTGTATAGTGTCTTCGCAATTGGAAAGATAAAGGAAACAAGGAACTCATCAAGAG 3642  
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058  
DB 3643 CTCTTGAGAAATCGAGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3696

AAD03506  
 ID AAD03506 standard; cDNA; 4279 BP.  
 XX  
 AC AAD03506;  
 XX  
 DT 13-JUN-2001 (first entry)  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.  
 XX  
 DE Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.  
 XX  
 OS Canis familiaris.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 17..3862  
 FT /tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype D) protein"  
 FT replace (91, T)  
 FT /tag= b  
 FT replace (607, C)  
 FT /tag= c  
 FT replace (1001, T)  
 FT /tag= c  
 FT replace (3458, A)  
 FT /tag= c  
 XX  
 PN W0200123540-A2.  
 XX  
 XX 05-APR-2001.  
 XX  
 XX 28-SEP-2000; 2000WO-US26767.  
 XX  
 XX 28-SEP-1999; 99US-0156510.  
 XX  
 XX (GENT-) GENTEST CORP.  
 XX  
 XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
 XX  
 DR WPI; 2001-235373/24.  
 DR P-PSDB; AAE00310.  
 XX  
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX  
 XX Claim 9; Page 102-107; 11pp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 XX  
 XX Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 8,38e-284 Length: 4279  
 Score: 3115.50 Matches: 602  
 Percent Similarity: 75.02% Conservative: 200

Best Local Similarity: 56.31% Mismatches: 220  
 Query Match: 58.45% Indels: 47  
 DB: 22 Gaps: 6  
 US-09-873-409-4 (1-1058) x AAD03506 (1-4279)  
 QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23  
 DB 734 TCATTTACTGATAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793  
 QY 24 LeuSerSerIleAArgThrValIleAlaPheAArgAlaGlnGluLysGluLeuGlnAArgSer 43  
 DB 794 TTAGCAGCAATCAGAACTGTGATTCCTTTGGAGGCAAAAGAAAGAACTTGAA----- 847  
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhePheGlnTrpLeuLeuSerCys 63  
 DB 847 ----- 847  
 QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
 DB 848 -----AGGTACAACAAAAATTTAGAAAGAACTAAAGGAATTTGGATAAAG 892  
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
 DB 893 AAAGCTATCACGCCCAACATTTCTATTGTGGCGCTTTCTATTGTATCTATGTCATCATAT 952  
 QY 104 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
 DB 953 GCTCTGGCTTTCTGGTATGGGACCTCTTGGTCTCTCCAGTGAA-----TATATATT 1006  
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
 DB 1007 GGACAAAGTACTACTCTCTTCTTCTGATTAAATGGGCTTTTAGTATTGGACAGGCA 1066  
 QY 144 ValProHisPheGluThrPheAlaIleAlaAArgGlyAlaAlaPheHisIlePheGlnVal 163  
 DB 1067 TCCCAAGCATTTGAAGCATTTTGCAAACGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1126  
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
 DB 1127 ATTGACAATAAACCAAGCATTTGACAGCTATTGCAAGAGTGGACATAAACAGATAATATT 1186  
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
 DB 1187 AAGCGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCTCTCGAAAAAGAGTTAAG 1246  
 QY 204 IleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
 DB 1247 ATCTTAAGGGTCTCAACCTGAAGGTTGAGAGTGGGACAGTGGCGCTGGTGGGAAC 1306  
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
 DB 1307 AGTGGCTGGCGGAAGAGCAGCAGCGTGCAGCTGATGCGAGGGCTCTATGACCCACAGAT 1366  
 QY 244 GlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnValArgHisTyrArgAsp 263  
 DB 1367 GGCATGCTGCTGATTGATGGACAGGACATTAGGACCAATAATGTAAGGCATCTTCGGAA 1426  
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
 DB 1427 ATTACTGGTGTGGTGTGAGTCAGGAGCCTGTGTTGTTTGGCCACCACGATAGTGAACATT 1486  
 QY 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaAArgGluAlaAsn 303  
 DB 1487 CGTATGGCCCGGAATAATGTCACCATGGATGAGATGAGAAAGCTGTTAAGGAAGCAAT 1546  
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
 DB 1547 GCCTATGATTTTATCATGAATACTACCTAATAAATTTGACACTCTGTTGGAGAGAGAGGG 1606  
 QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsn 343  
 DB 1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTTCGAAC 1666

Qy	344	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1667	CCCAAGATTCTTCTGCTGGATGAGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG	1726
Qy	364	ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	393
Db	1727	GTTTCAGGTGGCCCTGGATGAAGCCGAAAGCCGGAGCTACCATTTGATAGCTCATCTGT	1786
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1787	TTGTCTACAGTTGCTTAATCCGATGTCATTGCTGGTTTGATGATGGAGTCATTGTCTGAG	1846
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSer	423
Db	1847	AAAGGAAATCATGATGAATCTATGAAGAGAGAGGCGCATTTACTTCMAAACTTGTCACAATG	1906
Qy	424	GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTySerThrGluArgLys	443
Db	1907	CAG---ACAAGAGGAATGAAATGATGTAGAAAATGCCATGCTGTGAATCCAAAAGTGAA	1963
Qy	444	ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle-----	460
Db	1964	AGTGATGCCCTTGGAAATG---TCTCCAAAAGATTGAGGTCAGGTCAGTTTATAAAAAAGA	2020
Qy	461	-----AspLysAlaGluGluSerThrGln	468
Db	2021	TCAACTCGCAGGAGTATACATGCACCACAAAGCCAAAGACAGAAAGCTTGGTACAAAAGAG	2080
Qy	469	SerLysGluIleSerLeuProGluValSerLeuLysLysIleLeuLysLeuAsnLysPro	498
Db	2081	GACTTGAATGAGAAATGACTCCAGTTTCCCTTCTCGAGGATTCCTGAAGCTGAATCAACT	2140
Qy	489	GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	508
Db	2141	GAATGGCTTATTTTGGTGGTATATTTTGTGCTATTATAACGGAGGCGCTGCACCA	2200
Qy	509	ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr	527
Db	2201	GCATTTTCAATAATATTTCAAGGATTATAGGATCTTTTCCCGAGATGAGGATCCTGAA	2260
Qy	528	ThrLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCys	547
Db	2261	ACAAAACGACAGAAATAGTAACATGTTTCTGTATGTTCTTAGTCTCTTGGAAATTTTCT	2320
Qy	548	PheValSerTyrrPheMetGlnGlyLeuPheTyrrGlyArgAlaGlyGluIleLeuThrMet	567
Db	2321	TTTATTACATTTTCTCCAGGCTTCACATTTGGCAAGCTGGGAGATCCTCACTAAG	2380
Qy	568	ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrrGlnAspIleAlaTrpPheAspGlu	587
Db	2381	CGGCTTCGATACATGTTTTCAGATCCATGCTGACAGCAGGATGTCAGCTGGTTTGATC	2440
Qy	588	LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln	607
Db	2441	CCTAAAAACACCATCGAGCATTTGACAAACAGGCTTGGCAATGATGCGGCTCAAGTTAAA	2500
Qy	608	GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer	627
Db	2501	GGGCTATAGTTTCCAGGCTTGCTGTTCATCCCAAGATATAGCAAAATCTTTGGGACAGGC	2560
Qy	628	ValIleIleSerPheIleTyrrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro	647
Db	2561	ATTATTATATCCTTAATCATGTTGGCAATTAAACATTTTACTTCTTAGCAATTTGATCCC	2620
Qy	648	ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp	667
Db	2621	ATCATTTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGTCTGGCAAGACACTGAAAGAT	2680
Qy	668	LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr	687
Db	2681	AAGAAAGAGCTAGCAAGAGCTGGGAAGATTGTCACAGAAGCCATCGAAATCTCCGAAC	2740

Qy	588	IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr	707
Db	2741	GTTCCTTTCTTTGACTCGGAGCAGCAAGTTTGAATACATGTATGCACAGAGTTTCCAAAGTA	2800
Qy	708	GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis	727
Db	2801	CCATACAGAAACTCTTTGAGGAAGACACACATCTTCGGGGTCTCAATTTTCTATCACCCAG	2860
Qy	728	AlaPheIleTyrPheAlaIleAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla	747
Db	2861	GCAATGATGATATTTTCCATGCTGCGCTGTTTCGGSTTGGTGCCTACTTGGTGGCAAAAT	2920
Qy	748	GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet	767
Db	2921	GAGTTCATGAACCTTTCCAGGATGTTCTTTTGGTGATTTCTCAGCTATTGTCTTTGGTGCCATG	2980
Qy	768	AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla	787
Db	2981	GCAGTGGGCAGGTCAAGTTCATTTGCTCTGACTATGCAAAAGCCAAAGTATCAGCAGCC	3040
Qy	788	HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys	807
Db	3041	CACGTCAATCATGATCATTTGAAAAGAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC	3100
Qy	808	LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys	827
Db	3101	AAGCCAAATACGTTGGAAGAAATGTGCATTTAAATAGGTGCTGTTTCAACTATCCCACT	3160
Qy	828	ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal	847
Db	3161	CGACAGACATCCCGTGCTCGAGGGCTGAGCCTCGAGGTGAGAAAGGCCACAGCGCTG	3220
Qy	848	AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu	867
Db	3221	GCCCTCTAGGTAGCAGTGGCTGTGGGAAGACACAGATTGTTTCACCTCCTAGAGCGCTTC	3280
Qy	868	TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal	887
Db	3281	TATGACCCCTTGGCTGGTTCAGTGTCTAAATGTATGATGGCAAGAGATAAAGCACCTGAATGTC	3340
Qy	888	GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer	907
Db	3341	CAGTGGCTCCGAGCACACCTCGGGCATCGTGTCTCAGGAGCCCATCTCTGTTGACTGTCAGC	3400
Qy	908	IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys	927
Db	3401	ATTCCGAGAAACATTCCTATGAGAGCACACGCCGGTCTGATCATCATGAGAGATTGTG	3460
Qy	928	GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn	947
Db	3461	CAGGCAGCAAGGAGGCCAATACACCACTTTCATCGAGACACTCCCTGAGAAATACAAAC	3520
Qy	948	ThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle	967
Db	3521	ACCAGATAGGACACAAGAAGAACCCAGCTCTCTGTGGCGCAGAAACAGCGCATTTGCCATA	3580
Qy	968	AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu	987
Db	3581	GCTCGCGCTCTTGTATAGACAGCTCATATTTTTCCTTGGATGAAGCTACATCAGCTCTG	3640
Qy	988	AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr	1007
Db	3641	GATACAGAAAGTGAAAGGTTGTCCAAAGAGCCCTCGCAAAAGCCAGAGAAGGCGCCACC	3700
Qy	1008	CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu	1027
Db	3701	TGCATTGTGATTCGCCACCGCTGTGCCACCATCCAGATGTCAGATTTAATAGTGGTGT	3760
Qy	1028	HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle	1047
Db	3761	CAGAAATGGCAAAAGTCAAGGAGCATGGCACACATCAACAGCTCTCTGGCCAGAAAGGCATC	3820
Qy	1048	TyrPheLysLeuValAsnAlaGlnSer	1056



Qy	364	ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383
Db	1727	GTTCAAGTGGCCCTCGATTAAGCGCAGAAAGCCGCGACTACCATTTGTATAGCTCATCGT	1786
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1787	TTGTCTACAGTTCTGTAATGCCGATGTCATTTGCTGTTTGATGATGAGTCATTTGGAG	1846
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrtYrSerLeuValMetSer	423
Db	1847	AAAGGAAATCATGATGAATCTCATGAAGAGAGAGGCCATTTACTTTCAAACTTTGTCACAATG	1906
Qy	424	GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys	443
Db	1907	CAG---ACAAGAGAGAAATGAATTCGATTAGAAAATGCCCTGGTGAATCCAAAAGTGAA	1963
Qy	444	ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle-----	460
Db	1964	AGTAGTGCCTTGGAAATG---TCTCAAAGAAGTTCAGGGTCAGTTTAATAAAAAGAAGA	2020
Qy	461	-----AspLysAlaGluSerThrGln	468
Db	2021	TCAACTCGCAGGAGTATACATGCACCACAAGCCCAAGACAGAAAGCTTGGTACAAAAGAG	2080
Qy	469	SerLysGluLeSerLeuProGluValSerLeuLysLysIleLeuLysLeuAsnLysPro	488
Db	2081	GACTTGAATGAGAATGTACTCCAGTTTCCTTCTCGAGAGATTCCTCAAGCTGAACTCAACT	2140
Qy	489	GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	508
Db	2141	GAATGGCTTATTTTGTGGTGGTATATTTTGTGCTATATAACGGAGGCGCTGCAACCA	2200
Qy	509	ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr	527
Db	2201	GCATTTTCAATAATATTTTCAAGGATTA TAGGGATCTTTACCAGAGTAGGATCCTGAA	2260
Qy	528	ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys	547
Db	2261	ACAAAACGACAGAATAGTAACATGTTTCTGTATTGTTCTAGTCCTTGGAAATTAATTCT	2320
Qy	548	PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet	567
Db	2321	TTTATTACATTTTCTCCAGGCTTCACATTTGGCAAGCTGGGAGATCCTCACTAAG	2380
Qy	568	ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu	587
Db	2381	CGGCTTCGATACATGTTTTCAGATCCATGCTGACAGCAGGATGCAGCTGTTTGATGAC	2440
Qy	588	LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln	607
Db	2441	CCATAAAACACCACTGGAGCATTGCAACACGAGCTTGGCAATGATGGCGCTCAAGTTTAA	2500
Qy	608	GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer	627
Db	2501	GGGCTATAGTTTCCAGGCTTGCTGTCATTACCCAGATATAGCAATCTTGGGACAGGC	2560
Qy	628	ValIleLeSerPheIleTyrGlyTrpGluMetThrPheLeuLeuSerIleAlaPro	647
Db	2561	ATTATTATATCCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATTTGTACCC	2620
Qy	648	ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp	667
Db	2621	ATCAATTCGAATAGCAGAGTTGTTGAAATGAAAAATGTTTGTCTGGCAAGCACTGAAAGAT	2680
Qy	668	LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr	687
Db	2681	AAGAAAGAGCTAGNAGGAGCTGGAGAGATTGCTACAGAAGCATCTGAAAACTCCGAACT	2740
Qy	688	IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr	707
Db	2741	GTTCGTTCTTTTGACTCGGAGACAGAAGTTGAATACATGATGATGACAGAGTTTCAAGTA	2800

## RESULT 4





[illegible]

RESULT 5  
AAD03505  
ID AAD

AAD03505 standard; cDNA; 4279 BP.



XX  
AC AAD03505;  
XX  
XX 13-JUN-2001 (first entry)  
XX  
XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.  
XX  
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
KW MDR1; drug bioavailability; transgenic animal; genetic model; ss.  
XX  
XX Canis familiaris.  
OS  
XX  
XX Key Location/Qualifiers  
FH 17..3862  
FT CDS /tag= a  
FT /product= "dog P-glycoprotein (PGP) allelic variant  
FT (Genotype B) protein"  
FT allele replace (91, T)  
FT /tag= b  
FT allele replace (607, C)  
FT /tag= c  
XX WO200123540-A2.  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-US26767.  
XX  
XX 28-SEP-1999; 99US-0156510.  
XX  
XX (GENT-) GENTEST CORP.  
XX  
XX Stocker PJ, Steimel-crespi DT, Crespi Ch, Reif TC, Patten CJ;  
XX WPI; 2001-235373/24.  
DR P-PSDB; AAE00309.  
XX  
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT PGP inhibitors -  
XX  
XX Claim 9; Page 93-99; 11lpp; English.  
XX  
XX The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity, by  
CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump  
CC exporting small molecules across the cell membrane. This enzyme  
CC is a member of the ABC transporter family.  
XX  
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;  
SQ

Alignment Scores:  
Pred. No.: 1.61e-283 Length: 4279  
Score: 3112.50 Matches: 601  
Percent Similarity: 75.02% Conservative: 201  
Best Local Similarity: 56.22% Mismatches: 220  
Query Match: 58.40% Indels: 47  
DB: 22 Gaps: 6

US-09-873-409-4 (1-1058) x AAD03505 (1-4279)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23  
DB 734 TCATTACTGATAAAGAACTCTGGCCATGCAAAAGCTGGAGCTAGCTGAGAACTC 793  
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
DB 794 TTAGCAGCAATCAGAACTGTGATTGCCITTTGAGGACAAAGAAAGAACTTGAA-----847  
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
DB 847 -----847  
QY 64 ValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
DB 848 -----AGGTACAAACAAAAATTTAGAAAGCTAAAGAAATTTGGATAAAG 892  
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
DB 893 AAAGCTATCAGGCCAACATTTCTATTGGTGGCGCTTCTTATTGATCTCTATGCATCATAT 952  
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
DB 953 GCTCTGGCTTTCTGGTATGGGACCTCTTGGTCTCTCCAGTGAA-----TATTCTATT 1006  
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
DB 1007 GGACAAAGTACTCCTCTCTCTTTCTGTATTAAATTTGGGGCTTTTATGATTATGGACAGGCA 1066  
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
DB 1067 TCCCAAGCATTTGACATTTGCCAAGCGAAGAGGAGAGCTTATGCAATCTTCAAGATA 1126  
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
DB 1127 ATTGACAATAAACAAGCATTTGACAGCTATTTCGAAGATGGACATAAACACAGATAATATT 1186  
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
DB 1187 AGGGAAATTTGGAATTTCAAAATTTCTACCTTCAGTTACCTTCTCGAAAAGAAAGTTAAG 1246  
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
DB 1247 ATCTTAAGGGTCTCAACTGAAGTTTCAGATGGGACAGACAGTGGCGCTGGTGGGAAC 1306  
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
DB 1307 AGTGGCTGGGGAAAGACGACGCCGTGAGCTGATGACAGAGGCTCTATGATACCCACAGAT 1366  
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
DB 1367 GGCATGCTCTGTATTGATGGACAGACATTTAGGACCCATTAATTAAGGCATCTTCGGGAA 1426  
QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
DB 1427 ATTACTGGTGTGTGAGTTCAGGAGCCCTGTGTGTTTCCACACCATGATAGTGAACATT 1486  
QY 284 LysTyrGlyArgAspValThrAspGluLeuMetGluArgAlaAlaArgGluAlaAsn 303  
DB 1487 CGCTATGGCGCGAAATGTCCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCAAT 1546  
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323  
DB 1547 GCCTATGATTTTATCATGAAACTACCTTAATAATTTGACACTCTGGTTGGAGAGAGGG 1606  
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
DB 1607 GCCAGCTGAGTGTGGACAGAAACAGAGAAATGCCAATGCTCGGCCCTGTTGCGAAC 1666  
QY 344 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
DB 1667 CCAAGATTTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726  
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 383

Db	1727	GTTCAAGTGGCCCTGGATAAGGCCAGAAAGAGCCGGACTACCATTTGATAGTCATCGT	1786
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1787	TTGTCTACAGTTCTGTAATCCCATGTCATTTGCTGGTTTGGATGGAGTCATTTGGAG	1846
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTytyrSerLeuValMetSer	423
Db	1847	AAAGGAAATCATGATGAACTCATGAAGAGAGAGGCCATTTACTTCAAACTTGTCACAATG	1906
Qy	424	GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTytyrSerThrGluArgLys	443
Db	1907	CAG---ACAAGAGGAAATGAAATGAGTTAGTAAATGCCACTGGTGAATCCAAAGTGAA	1963
Qy	444	ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle-----	460
Db	1964	AGTGATGCTTGGAAATG---TCTCAAAGAGTTCAGGGTCCAGTTTAATAAAGAGAGA	2020
Qy	461	-----AspLysAlaGluGluSerThrGln	468
Db	2021	TCAACTCGCAGGAGTATACATGCACCAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG	2080
Qy	469	SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro	488
Db	2081	GACTTGAATGAGAAATGACTCTCCAGTTTCTTCTCGAGGATTCCTGAAGCTGAATCAACT	2140
Qy	489	GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	508
Db	2141	GAATGGCTTATTTTGTGGTGGTATATTTGTGCTATATAACGGAGGGCTGCAACCA	2200
Qy	509	ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr	527
Db	2201	GCATTTTCAATAATATTTTCAAGGATATAGGATCTTTTACCAGAGATGAGGACTCTGAA	2260
Qy	528	ThrLeuLysHisAspAlaGluIleTytyrSerMetIlePheValIleLeuGlyValIleCys	547
Db	2261	ACAAAACGACAGAATAGTAACATGTTTCTGCTATTGTTTCTAGTCTCTTGGAAATTTATTTCT	2320
Qy	548	PheValSerTytyrPheMetGlnGlyLeuPheTytyrGlyArgAlaGlyGluIleLeuThrMet	567
Db	2321	TTTATTACATTTTCTCCAGGCTTCATTTGGCAAGCTGGGAGATCCTCACTAAG	2380
Qy	568	ArgLeuArgHisLeuAlaPheLysAlaMetLeuTytyrGlnAspIleAlaTrpPheAspGlu	587
Db	2381	CGCTTCGATACATGTTTTCAGATCCATGCTGACAGAGGATGCAGCTGGTTTGTATGAC	2440
Qy	588	LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln	607
Db	2441	CTTAATAAACACCACTGGAGCATTTGACACCGGCTTGGCAATGATCGGCTCAAGTTTAAA	2500
Qy	608	GlyValThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer	627
Db	2501	GGGGCTATAGTTTCCAGGCTTGCTGTCAATCCCAAGATATAGCAAAATCTTGGGACAGGC	2560
Qy	628	ValIleIleSerPheIleTytyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro	647
Db	2561	ATTATTATATCTTATCTATGTTGGCAATTAACATTTTACTTCTTAGCAATTTGATCCC	2620
Qy	648	ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp	667
Db	2621	ATCATTTGCAATAGCAGGAGTTGTTGAAATGAAATGTTGTCTGGCAGACGACTGAAAGAT	2680
Qy	668	LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr	687
Db	2681	AGAAGAGAGCTAGAAGAGCTGGGAAGATTGCTACAGAAGCCATCGAAATCTCCGAAC	2740
Qy	688	IleValSerLeuThrArgGluLysAlaPheGluGlnMetTytyrGluGluMetLeuGlnThr	707
Db	2741	GTGTGTTCTTTGACTCGGAGCAGAGTTTGAATACATGATGACACAGAGTTTGCAGTA	2800
Qy	708	GlnHisArgAsnThrSerTytyrLysAlaGlnIleIleGlySerCysTytyrAlaPheSerHis	727

RESULT 6  
AAD03489  
ID AAD03489 standard; CDNA: 4317 BP.



Db 1837 TTGCTACAGTTCTGTAATCCCGATGTCATTTGCTGGTTCATGATGAGTCAATGTCGAG 1896  
Qy 404 LysGlyValHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer 423  
Db 1897 AAGGAAATCATGATGAATCATGAAGAAGAGGGGATTTACTTCAAACTTGTCACAATG 1956  
Qy 424 GinAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443  
Db 1957 CAG---ACAAGAGGAATGAAATGAGTTAGAAATGCCACTGGTGAATCCAAAGTGA 2013  
Qy 444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle----- 460  
Db 2014 ACTGATGCCTTGGAAATG---TCTCCANAAGATTCCAGGCTCCAGTTTAATAAAGAA 2070  
Qy 461 -----AspLysAlaGluGlnSerThrGln 468  
Db 2071 TCNACTCGCAGGAGTATACATGCACCACCAAGGCCAGACAGAAGCTTGGTACAAAAGAG 2130  
Qy 469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488  
Db 2131 GACTTCAATGAGAATGTTCTTTCAGTTTCTCTGAGGAGTTCTGAAGCTGAACCTCAA 2190  
Qy 489 GluTrpPropheValValLeuGlnGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508  
Db 2191 GAATGGCCTTATTTTCTGGTGGTATATTTTGTGCTATTATAAAGGAGGCTTCGAACCA 2250  
Qy 509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527  
Db 2251 GCATTTTCAATATATTTTCAAGGATTTATAGGATCTTTACCCGAGATGAGGATCTCGAA 2310  
Qy 528 ThrLeuLysHisAspAlaGluIleIleSerMetIlePheValIleLeuGlyValIleCys 547  
Db 2311 ACRAAAGCAGACAGATAGTAACATGTTTCTGTATTGTTCTAGTCTTGGAAATTATTTCT 2370  
Qy 548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567  
Db 2371 TTTATTACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCTCTCACTAAG 2430  
Qy 568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587  
Db 2431 CGGCTTCGATACATGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGC 2490  
Qy 588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607  
Db 2491 CCTAAAACACCACTGGAGCATTCACACACAGGCTTCCCAATGATCGGCTCAAGTTAA 2550  
Qy 608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627  
Db 2551 GGGGCTATAGGTTCCAGGCTTGCTGTCTATTCACCAAGATATAGCAAAATCTTGGGACAGGC 2610  
Qy 628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647  
Db 2611 ATTATTATATCTTAACTAATGTTGGCAATTAACATTTTACTCTTAGCAATGTACCC 2670  
Qy 648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667  
Db 2671 ATCATTTGCAATACAGAGGATGTTGTAATGAAATGTTGTCGGACAGCACTCAAGAT 2730  
Qy 668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687  
Db 2731 AAGAAAGAGCTAAGAGGAGCTGGGAAGATTGCTACAGAGCCATCGAAAACCTTCCGAAC 2790  
Qy 688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707  
Db 2791 GTTGTTCCTTGTGATCGGAGCAGAGTTTGAATACATGATGCACAGAGTTTGCAGTA 2850  
Qy 708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727  
Db 2851 CCATACAGAAACTCTTTGAGAAAGCACACATCTTCGGGGTCTCAATTTCTATCACCAG 2910  
Qy 728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747  
Db 2911 GCAATGATGATTTTCTCTATGCTGGCTGTTTCCGGTTTGGTGCTACTTGTGGTGAAT 2970

Qy 748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767  
Db 2971 GAGTTTCATGAACCTTTCAGGATGTTCTTTGGTATTCTCAGCTATTGCTCTTGGTGCATG 3030  
Qy 768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787  
Db 3031 CGAGTGGGCGAGCTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAAGATATCAGCAGCC 3090  
Qy 788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLys 807  
Db 3091 CACGTCATCATCATCATTTGAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAGGCGCTC 3150  
Qy 808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827  
Db 3151 AAGCCAAATACGTTGGAAGAAATGTACATTTAATGAGTCTGTTCAACTATATCCACT 3210  
Qy 828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847  
Db 3211 CGACCAAGACATCCCGTGTCTCAGGGGCTGAGCCTCGAGTGAAGAGGGCCAGACGCTG 3270  
Qy 848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867  
Db 3271 GCCTCTGATAGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTTCAGCTCTCAGAGCGCTC 3330  
Qy 868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887  
Db 3331 TATGACCCCTTGGCTGGTTCACTGCTGCTAATTGATGGCAAGAGATAAAGCACCTGAATGTC 3390  
Qy 888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907  
Db 3391 CAGTGGCTCCGAGCACACCTGGGCACTGCTGTCTCAGGAGCCCATCTCTTTGACTCGAGC 3450  
Qy 908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927  
Db 3451 ATTCGCGAAGCAATTCCTTATGGAGACAACAGCCGGTCTGTATCATGATGAAGAGATTATG 3510  
Qy 928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947  
Db 3511 CAGGCAAGCAAGGAGGCCAACATACACCACTTCATCGAGACACTCTCCCTGAGAAATACAAC 3570  
Qy 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967  
Db 3571 ACCAGAGTAGGAGACAAGAAACCCAGCTCTCTGTGGCCAGAAACAGCGCATTTGCCATA 3630  
Qy 968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987  
Db 3631 GCTCGGCTCTTGTATGACAGCCTCATATTTTCTTTTGGATGAAGCTACATCAGCTCTG 3690  
Qy 988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007  
Db 3691 GATACAGAAAGTGAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGCCGCCACC 3750  
Qy 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027  
Db 3751 TGCATTGTGATGCCCGCTTGTCCACCATCCAGATCCAGATTTAATAGTGGTGT 3810  
Qy 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAsnArgAspIle 1047  
Db 3811 CAGAATGGCAAGTCAAGGAGCATGGCACATCAACAGCTGCTGCTCAGAAAGGCATC 3870  
Qy 1048 TyrPheLysLeuValAsnAlaGlnSer 1056  
Db 3871 TATTTTCCATGATGATGTCAGGCT 3897  
RESULT 7  
AAZ52047  
ID AAZ52047 standard; cDNA; 4369 BP.  
XX  
AC AAZ52047;  
XX  
DT 18-JUL-2000 (first entry)  
XX

DE Rat multidrug resistance protein 1a cDNA.

XX Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;  
 KW drug formulation; formulation excipient; compound design; inflammation;  
 KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;  
 KW central nervous system disorder; auto-immune disease; kidney disease; ss.  
 XX Rattus rattus.

XX Key Location/Qualifiers  
 FT CDS 352..4170  
 FT /tags= a  
 FT /product= "Rat multidrug resistance protein"

XX W0200015650-A1.  
 XX 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20770.  
 XX 17-SEP-1998; 98US-0156800.  
 XX 09-DEC-1998; 98US-0208809.  
 XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;  
 XX WPI; 2000-271372/23.  
 XX P-PSDB; AAY70596.

XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays  
 PT to provide information on drug formulation, selection of formulation  
 PT excipients and compound design

XX Claim 2; Page 23-24; 33pp; English.

XX The present cDNA sequence encodes rat multidrug resistance  
 CC protein 1a (mdrla). This protein is a member of the multi-specific drug  
 CC transporters family. Mdrla is used in assays to provide information on  
 CC drug formulation, selection of formulation excipients and compound  
 CC design. They are used in cell based, membrane based, binding or other  
 CC assays to provide information that may enhance drug formulation. This  
 CC invention further relates to the generation of in vivo and in vitro  
 CC comparison data to predict oral absorption and pharmacokinetics. This  
 CC enables the selection of drugs with optimal pharmacokinetics, i.e. good  
 CC oral bioavailability, brain penetration, plasma half life, and minimum  
 CC drug interaction. Transgenic and knock-out animals created using rat  
 CC mdrla provides an insight into treating and preventing human diseases  
 CC including cancer, inflammation, cardiovascular disease, central nervous  
 CC system disorders, auto-immune and kidney disease.

XX Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:  
 Pred. No.: 6,13e-283 Length: 4369  
 Score: 3106.50 Matches: 600  
 Percent Similarity: 75.54% Conservative: 206  
 Best Local Similarity: 56.23% Mismatches: 218  
 Query Match: 58.28% Indels: 43  
 DB: 21 Gaps: 7

US-09-873-409-4 (1-1058) x AA252047 (1-4369)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23  
 DB 1039 TCATTTTACTATAAGGAACCTCCAGCTTATGCAAAAGCTGAGCAGTGTGCTGAAGAAGTC 1098  
 QY 24 LeuSerSerLysAlaGlnAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
 DB 1099 TTAGACCCATCAGAACTGTGATTGCTTTGGAGGACAAAGAGGAACTTGAA----- 1152  
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db 1152 ----- 1152

QY 64 ValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
 DB 1153 -----AGGTACAATAACATTTTGAAGAAGCTAAAGAGCTTGGATAAAG 1197

QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
 DB 1198 AAAGCTATCACGGCCAAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATCATAT 1257

QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123  
 DB 1258 GCTCTGGCATTCGGTATGGGACTTCCTTGGTCATCTCAAAAGAA-----TACACTATT 1311

QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143  
 DB 1312 GGACAAGTCTCACTGCTCTTTTCTGTATTAAATGGAGCATTCAGTGTGGCAGGCA 1371

QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
 DB 1372 TCTCCAATATTGAAGCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTTCAGTATA 1431

QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
 DB 1432 ATTGATAATAAGCCCACTATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACACATA 1491

QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
 DB 1492 CAAGGAATTTGGAATTCAAAATATTCTACTTCAGTTACCCGCTCGAAAAGACGTTGAG 1551

QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
 DB 1552 ATCTTGAAGGGCTCAACCTGAAGGTGAAGCGGCGCAGCGTAGCCCTGTTGGCAAC 1611

QY 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
 DB 1612 AGTGGCTGTGGGAAAGCAACACTGTCCAGCTGCTCCAGAGGCTCTACGACCCCATAGAG 1671

QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
 DB 1672 GCGAGGTGAGTATCCGCGGACAGGACATCAGACCATCAATGTGAGGTATCTGCGGGAA 1731

QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
 DB 1732 ATCATTTGGGTGTGAGTCAGGAACCGTCTGTTGCCACCAACAATTGCGGAAACATT 1791

QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
 DB 1792 CGCTATGGCGGAGAAACGTCACCATGGATGATAGAGAAAGCTGTCAAGGAGGCAAT 1851

QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
 DB 1852 GCCTATGATTTTCATGTAAGAACTGCCCCCAAAATTTGACACCCCTGTTGGTGAGAGGG 1911

QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
 DB 1912 GCGCAGCTGAGTGGGCGACAGAAACAGAGGATCGCCATTCGCCGGGCCCTGTGTCGCAAC 1971

QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
 DB 1972 CCCAAGATCCCTTTTGTGGATGAGGCCACGTCAGCTTGGACACAGAAAGCGGAGCGGTG 2031

QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
 DB 2032 GTTCAGCCCTCTGGATAGGCTAGAGAGCGCGGACCAACCATTTGTAGTAGTCCACGC 2091

QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrIleLysAspGlyMetLeuAlaGlu 403  
 DB 2092 TTGTCTACAGTTCGCAATGCTGACGTCATTTGCTGTTTTCATGGTGTGTCATTTGTGAG 2151

QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
 DB 2152 CAAGGAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTCAAACTTGTGATGACT 2211

```
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 435
Db 2212 CAGACAGCAGGAAATGAAATTCGAATTAGGAATGAAGCTTGTGAATCTAAGACGCGAATT 2271
Qy 436 -----MetThrTySerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 2272 GATAATGTGCAGCATGCTCTCAAAAGATTCCGGGATCCAGTCTAATAAGAAAGAGATCAACT 2331
Qy 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db 2332 CGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACCTTAGCACCAAGAGGCTCTG 2391
Qy 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTTP 490
Db 2392 GATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 2452 CCTATTATTGTTGGTGTATTCTTGTGCNTAATAAATGGAGGCTTGCACCCAGCATTC 2511
Qy 511 SerIlePheAlaLysIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
Db 2512 TCATANTATTTCAAAGGTGTAGGGTTTTCACAAAATGACACCCCTGNAATCCAG 2571
Qy 530 LysHisAspAlaGluIleThrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2572 CGGCAGAACAGCAACTGTTTCTTATTGTTCTGATCCTTGGGATCATCTCTTTCAAT 2631
Qy 550 SerTyPheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2632 ACGTTTTTCTTCAGGGCTTCACATTGGCAAGCTGGAGAGATCCTCACCAAGCGACTC 2691
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyArgAspIleAlaThrPheAspGluLysGlu 589
Db 2692 CGATACATGCTTCAAAATCCATGCTGAGACAGACATAGCTGGTTGATGACCCCTAA 2751
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2752 AACACACAGGCGCTGACCACGAGCTTGCCCAATGACGCTCTCAAGTGAAGGGCT 2811
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2812 ACAGGGTCTAGGCTGCTGTTATTACCAACATAGCAAAATCTGGGACAGCATCATC 2871
Qy 630 IleSerPheIleTyArgLysGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2872 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCTCTAGCAATTTGTTCCCATCATT 2931
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db 2932 GCTATAGCAGGAGTGTGAAATGAAATGTTCTCTGGACAAGCGCTGAAGATAGAAG 2991
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db 2992 GAACTAGAAGTTCTGGGAAGATCGCTACAGAAGCAATTGAAACITTTCCACTGCTGCTC 3051
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeuGlnThrGlnHis 709
Db 3052 TCTTTGACTCGGAGCAGAAAGTTTGAACACTATGTATGTCGCCAGAGCTTGCAGATACCATAC 3111
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyArgAlaPheSerHisAlaPhe 729
Db 3112 AGAAATGCTTTGAAGAAAGCGACGCTTTGGGATCATCTTCTCTTCCACCCAGGCCATG 3171
Qy 730 IleTyPheAlaTyArgAlaAlaGlyPheArgPheGlyAlaTyLeuIleGlnAlaGlyArg 749
Db 3172 ATGTATTCTCTCTGCTGCTGTTTCCGGTTCATGCTCTACTTGGTGGCAGCAGAACTC 3231
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyArgLysAlaMetAlaIle 769
Db 3232 ATGACATTTGAAAAATGTTCTGTAGTATTCTCAGCTATTCTTGTGTGCTGCGCAGTG 3291
```

```
Qy 770 GlyLysThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 3292 GGGCAGGTCAAGTTCATTCGCTCTGACTACGCGAAGCAAAAGTCTCGCATCCACATC 3351
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3352 ATCAGATCATTTGAGAAATCCCTGAGATTGACAGCTACAGCACGGAGGCTTGAAGCT 3411
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyProCysArgPro 829
Db 3412 AATATGTTGGAAGGAAATGTAATTTAATGAGTTCATGTTCAACTATCCACCCGACCC 3471
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3472 AACATCCCACTGCTTCAGGGGCTGAGCTTAGAGTGAAGAAAGGGCAGACGCTGGCCCTC 3531
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyArgP 869
Db 3532 GTGGCAGCAGTGGCTGCGGGAAGAGTACAGTGTCTCAGCTGCTTGAGGCTTCTATGAC 3591
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTTP 889
Db 3592 CCCATGGCCGGAACAGTGTTCCTAGATGGCAAAAGAAATAAGCAACTCAATGTCCAGTGG 3651
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3652 CTCGCGCCCACTGGGCATTTGTCTCCAGAGGACCCATCTCTGTTGACTGCAGCATCGCC 3711
Qy 910 GluAsnIleAlaTyArgLysAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db 3712 GAGAACATTCCTACGGAGACCAACAGCCGTGCTGTCTCTAATAAGAGATCGTGAAGGCA 3771
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgThrGln 949
Db 3772 GCCAAGGAGGCCAACATCCACAGTTCATCGACTCATCGCTGAGAAATACACACACAGA 3831
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3832 GTGGGAGACAAAGGGACTCAGCTGTGCGCGGGCAGAAAGCAGCGCATCGCCATCGCGCGC 3891
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3892 GCCTCGTCAGACAGCTCCTCACATCTTACTTCTGGATGAAGCCGACATCAGCTCTGGATACG 3951
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3952 GAGAGTGAAGAGTCTGTCAGGAAGCGCTGGACAAAGCCAGGGAAGCCGACCTGCATT 4011
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4012 GTGATCGCGCAGCGCTGTCCACCATCCAGAACCGCAGACTTGTATCGTGTGTATTCAGAAC 4071
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyPhe 1049
Db 4072 GGCAGGTCAAGGACGACCGGCCACCCACAGACTTGTATCGTGTGTATTCAGAAC 4131
Qy 1050 LysLeuValAlaGlnSer 1056
Db 4132 TCGATGCTCAGTGTGCGGCT 4152
RESULT 8
AAZ52048
ID AAZ52048 standard; cDNA; 4425 BP.
XX
AC AAZ52048;
XX
XX 18-JUL-2000 (first entry)
DE Rat multidrug resistance protein 1a cDNA derived from EST sequences.
XX Multidrug resistance protein 1a; mdrla; multi-specific drug transporter;
KW drug formulation; formulation excipient; compound design; inflammation;
KW pharmacokinetic; oral absorption; cancer; cardiovascular disease;
```



KW central nervous system disorder; auto-immune disease; kidney disease;  
 KW EST; expressed sequence tag; ss.

OS Rattus rattus.

XX Key Location/Qualifiers  
 FH 352..4170  
 CDS /\*tag= a  
 FT /product= "Rat multidrug resistance protein"

XX WO200015650-A1.

PN 23-MAR-2000.

XX 10-SEP-1999; 99WO-US20770.

XX 17-SEP-1998; 98US-0156800.

PR 09-DEC-1998; 98US-0208809.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;

XX WPI; 2000-271372/23.

DR P-PSDB; AAY70597.

XX Isolated rat mdrla polynucleotides and polypeptides, useful in assays  
 PT to provide information on drug formulation, selection of formulation  
 PT excipients and compound design -

XX Claim 11; Page 27-29; 33pp; English.

PS The present cDNA sequence encodes rat multidrug resistance protein 1a  
 CC (mdrla). This cDNA is derived from EST (expressed sequence tag)  
 CC sequences. Mdrla is used in assays to provide information on  
 CC drug formulation, selection of formulation excipients and compound  
 CC design. They are used in cell based, membrane based, binding or other  
 CC assays to provide information that may enhance drug formulation. This  
 CC invention further relates to the generation of in vivo and in vitro  
 CC comparison data to predict oral absorption and pharmacokinetics. This  
 CC enables the selection of drugs with optimal pharmacokinetics, i.e. good  
 CC oral bioavailability, brain penetration, plasma half life, and minimum  
 CC drug interaction. Transgenic and knock-out animals created using rat  
 CC mdrla provides an insight into treating and preventing human diseases  
 CC including cancer, inflammation, cardiovascular disease, central nervous  
 CC system disorders, auto-immune and kidney disease.

XX SQ Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:

Pred. No.: 6,25e-283 Length: 4425  
 Score: 3106.50 Matches: 600  
 Percent Similarity: 75.54% Conservative: 206  
 Best Local Similarity: 56.23% Mismatches: 218  
 Query Match: 58.28% Indels: 43  
 DB: 21 Gaps: 7

US-09-873-409-4 (1-1058) x AA252048 (1-4425)

QY 4 SerLeuThrSerLySgLuLeuSerAlaTySerLySgAlaGlyAlaValAlaGluVal 23  
 Db 1039 TCATTACTGATAGGAACCTCCAGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 1098  
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLySgLuLeuGlnArgSer 43  
 Db 1099 TTACGACCATCAGAACTGTGATTCCTTTGGAGGACAAAAGAGGAACCTTGAA----- 1152  
 QY 44 PheLeuLeuAsnIleThrArgTyAlaTrpPheTyPheProGlnTrpLeuLeuSerCys 63  
 Db 1152 -----AspIleLySgAlaAspGluClnMetGluSer----- 435  
 QY 64 ValLeu\*\*\*PheValArgTyThrGlnAsnLeuLySgAlaLySgPheGlyIleLyS 83

Db 1153 -----AGGTACAAATAACAATTTGGAGAAGCTAAAGAGGCTGGGATAAAG 1197  
 QY 84 ArgThrIleAlaSerLySgValSerLeuGlyAlaValTyPhePheMetAsnGlyThrTy 103  
 Db 1198 AAAGCTATCAGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTATCTATGATCATAT 1257  
 QY 104 GlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsnGlyGluProGlyTyThrIle 123  
 Db 1258 GCTCTGGCATTCCTGGTATGGGACTTCTTGGTCACTCAAAAGAA-----TACACTATT 1311  
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyTyCysIleGlyAlaAla 143  
 Db 1312 GGACAAAGTGTCTCACTGTCTTTTCTGTTAATTTGGAGCATTCAGTGTGGCAGGCA 1371  
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
 Db 1372 TCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431  
 QY 164 IleAspLySgProSerIleAspAsnPheSerThrAlaGlyTyLySgProGlySerIle 183  
 Db 1432 ATTGATAATAAGCCAGTATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACACATA 1491  
 QY 184 GluGlyThrValGluPheLySgAsnValSerPheAsnTyTyProSerArgProSerIleLyS 203  
 Db 1492 CAAGGAAATTTGGAAATTCAAAATATTCACTTCAGTTACCCGCTCGAAAAGACGTTCA 1551  
 QY 204 IleLeuLySgGlyLeuAsnLeuArgIleLySgSerGlyGluThrValAlaLeuValGlyLeu 223  
 Db 1552 ATCTTGAGGGCTCAACCTGAAGGTGAAGCGGCGAGAGCGGTAGCCCTGGTTGGCAAC 1611  
 QY 224 AsnGlySerGlyLySgSerThrValValGlnLeuLeuGlnArgLeuTyTyAspProAspAsp 243  
 Db 1612 AGTGGCTGTGGGAAAAGCACAACTGTCCAGTGTCTCAGAGGCTCTACGACCCCATAGAG 1671  
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp 263  
 Db 1672 GGCAGGTCAAGTATCGCAGCAGACGATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1731  
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
 Db 1732 ATATTGGGTGTGAGTCAAGAACCCGCTGTGTTGCCACCAATTCGCGAAACATT 1791  
 QY 284 LyTyTyGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
 Db 1792 CGCTATGGCCGAGAAACGTCACCATGGATGATAGAGAAGAGCTGTCAAGGAAGCCAA 1851  
 QY 304 AlaTyAspPheIleMetGluPheProAsnLeuPheAsnThrLeuValGlyGlyLySg 323  
 Db 1852 GCCTATGATTTCATGAAACTGCCCCCAAAATTTGACACCCCTGGTGTGAGAGAGG 1911  
 QY 324 AlaGlnMetSerGlyGlyGlnLySgGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
 Db 1912 GCGAGCTGAGTGGGGACAGAAACAGAGATCGCAATTCGCCGGCCCTGGTCCGCAAC 1971  
 QY 344 ProLySgIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLySgAla 363  
 Db 1972 CCCAAGATCCTTTGTTGATGAGCCACGTCAGCTTGGACACAGAAAGCGGCGTG 2031  
 QY 364 ValGlnAlaAlaLeuGluLySgAlaSerLySgTyArgThrThrIleValValAlaHisArg 383  
 Db 2032 GTTCAGGCCCTCTGGATAGGCTAGAGAGGCGGACCACTTGTGATAGTCCACCG 2091  
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLySgAspGlyMetLeuAlaGlu 403  
 Db 2092 TTGCTACAGTTCGCAATGCTGAUGTCATTCGTTTGTGATGGTGTGATGTGGAG 2151  
 QY 404 LySgGlyAlaHisAlaGluLeuMetAlaAlaArgGlyLeuTyTySerLeuValMetSer 423  
 Db 2152 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCACTACT 2211  
 QY 424 Gln-----AspIleLySgAlaAspGluClnMetGluSer----- 435  
 Db 2212 CAGACAGCAGGAAATGAAATTGAATTAGGAATGAAGCTTTGTGAATCTAAAGAUGGAATT 2271



```
QY 436 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451
Db 2272 GATAAATGTGGACATGCTTCAAAAGATTCRCGGATCCAGTCTAATAAGAAAGAACACT 2331
QY 452 ValLysSerIleLysSerAspPheIleAapLysAlaGluLysSerThr---GlnSerLys 470
Db 2332 CGCAAAAGCATTCGTGGGCCACATGATCAAGACGGGGAACCTTAGCACCAAGAGGCTCTG 2391
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAnLysPProGluTir 490
Db 2392 GATGACGACGTACTCCAGCTTCTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAnGlyThrValHisProValPhe 510
Db 2452 CCTATTATTTGGTGGTGTATTTGTGCCATAATAAATGAGGCTTGCACACGACATTC 2511
QY 511 SerIlePheAlaLysIleIleThrMetPheGlyAsnAsnAapLysThrThrLeu--- 529
Db 2512 TCCATAATATTTTCAAGGTTGTAGGGTTTTCACAAAATAACACCCCTGAAATCCAG 2571
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db 2572 CGCACAACAGCAACTGTGTTCTTATTGTTCTGATCCTTGGGATCATCTCTTCAAT 2631
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db 2632 ACCTTTTCTCTCAGGCTTCACATTTGGCAAGCTGGAGAGATCTCTCAACAGGCACTC 2691
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirPheAspGluLysGlu 589
Db 2692 CGATACATGCTCTTCAATCCATGCTGAGACAGGACATAGAGCTGGTTGATGCCCTAAA 2751
QY 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db 2752 AACCCACAGGACGCTGACCAACAGGCTTGCCCAATGACCTGCTCAAGTGAAGGGGCT 2811
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db 2812 ACAGGCTCTAGGCTTGCTGTTATTACCAGAACATAGCAAAATCTTGGGACAGGCATCATC 2871
QY 630 IleSerPheIleTyrGlyTirPleuMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db 2872 ATATCCCTGATCTACGCGTGGCAATGGACACTTTTACTCTAGCAATTTGTTCCCAT 2931
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaLeuLysAspLysGln 669
Db 2932 GCTATACGAGGAGTGGTTGAATGAAATGTTGTCGACACAGCGCTGAAGATAGAAG 2991
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAnIleArgThrIleVal 689
Db 2992 GAATAGAGGTTCTGGAGAGATCGCTACAGAGCAATTCGAAACTTTCGCACTGTGCTC 3051
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db 3052 TCTTTGACTCGGAGCAGAAAGTTTGAACATATGATGCCACAGCGCTTGCAGATACCATAC 3111
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db 3112 AGAAATGCTTTGAAGAAAGCGCAGCTTTGGGATTCATCTTCTCTCTTCAACCCAGGCGATG 3171
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleLeuIleGlnAlaGlyArg 749
Db 3172 ATGTAATTCCTATGCTGCTGTTGTCGGTTTCATGCTACTTGTGGGACGAGAACTC 3231
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db 3232 ATGACATTTGAAATGTTCTGTAGTATTCTCAGTATTGCTTTGCTGGTGCATGGCAGTG 3291
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db 3292 GGGCAGGTGATGTTCAATTCGCTCTGACTACGCGAAGCCAAAGTCTCGGCATCCCATC 3351
```

```
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db 3352 ATCAGGATCATTTAGAAAAATCCCTGAGATTGACAGCTACAGCAGGAGGCTTTGAAGCCT 3411
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db 3412 AATATGTTGGAAGAAATGTGAATTTAATGAGTTCATGTTCACTATCCACCCGACCC 3471
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db 3472 AACATCCAGTGTCTCAGGGCTGAGCTTAGAGGTGAAGAAAGGCGAGACGCTGGCCCTC 3531
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3532 GTGGGACAGAGTGGCTGCGGAGAGTACAGTGGTCCAGCTGCTTGAGGCTTCTATGAC 3591
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAnValGlnTir 889
Db 3592 CCATGCGCGGAACAGTGTCTTAGATGGCAAGAATAAAGCAACTCAATGTCCAGTGG 3651
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3652 CTCGCGCCACCTGGGCTTGTGCCAGGAGCCCATCTCTGTTGACTGCAGCATGCC 3711
QY 910 GluAnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db 3712 GAGAACATTCCTTACGAGACACAGCCGCTGCTCTCATAGGAGATCGTGAAGGCA 3771
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3772 GCCAAGGAGCCCAACATCCACAGTTCATCGACTCACTGCTGAGAAATACAAACCCAGA 3831
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleArg 969
Db 3832 GTGGGAGACAAAGGACTCAGCTGTCGGCGGCGCAGAGCGCATCGCCATCGCGCGC 3891
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3892 GCCTCGTCAGACAGCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3952 GAGAGTGAAGAGTCTGACGAAGCGCTGGCAAGCGCGGAGGAGGAGCCGACCTGCATT 4011
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4012 GTGATCGCGCACCGCTCTCCACCATCCAGACGCGAGACTTGATCGTGTGATTCAGAAC 4071
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4072 GGCAGGTCAAGGAGCAGCGCACCCACAGCAGCTGTGCGCCCAAGAGGATCTATTTC 4131
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 4132 TCGATGCTCAGTGTGCGGCT 4152
RESULT 9
AAZ49332
ID AAZ49332 standard; cDNA; 3860 BP.
XX
AC AAZ49332;
XX
DT
XX 14-MAR-2000 (first entry)
XX
DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW cytokine; wild-type; ds.
XX
```

OS Homo sapiens.  
 XX Location/Qualifiers  
 FH 1..3843  
 FT /tags: a  
 FT /product= "Human wild-type MDR-1 protein"  
 FT replace (553..555, GTT)  
 FT /tag: b  
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given  
 XX in AAZ49333"  
 PN WO961589-A2.  
 PD 02-DEC-1999.  
 XX 27-MAY-1999; 99WO-US11825.  
 XX 28-MAY-1998; 98US-0086988.  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Sorrentino B, Bunting K;  
 DR WPI; 2000-072615/06.  
 DR P-PSDB; AAY58186.  
 XX Ex vivo expansion of hematopoietic stem cells transduced with a  
 PT sequence encoding human multidrug resistance-1, used for bone marrow  
 PT transplantation -  
 XX Claim 10; Page 68-70; 113pp; English.  
 CC This sequence represents cDNA encoding human wild-type  
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane  
 CC efflux pump, responsible for the export of drugs from cells,  
 CC particularly cancer cells. Wild-type MDR-1 shows increased  
 CC resistance to etoposide and decreased resistance to vinca  
 CC alkaloids compared with a mutant form (AAY58187) where the Gly at  
 CC position 185 is replaced by Val. The invention relates to transducing  
 CC hematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified hematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC hematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from hematopoietic stem cells, e.g., thalassemia,  
 CC Gaucher's disease, sickle cell anemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in hematopoietic stem cells.  
 CC Hematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC hematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.  
 XX  
 SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,88e-282 Length: 3860  
 Score: 3100.50 Matches: 595  
 Percent Similarity: 75.54% Conservative: 211  
 Best Local Similarity: 55.76% Mismatches: 218  
 Query Match: 58.17% Indels: 43  
 DB: 21 Gaps: 6  
 US-09-873-409-4 (1-1058) x AAZ49332 (1-3860)  
 QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23  
 Db 712 TCATTACTGATAAAGAACTCTTAGCTGTCACAAAGCTGAGCAGTAGCTGAAGAGTC 771  
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db 772 TTGGCAGCAATTAGAACTGTGATTGCTATTGGAGGACAAAGAAAGAACTTGAA----- 825  
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
 Db 825 ----- 825  
 QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
 Db 826 -----AGGTACACACAAATTTAGAGAGAGCTAAAGAAATGGGATAAG 870  
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
 Db 871 AAAGCTATTACAGCCAATATTCTATAGTGTCTCTTCTCTGCTGATCTATGATCTTAT 930  
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
 Db 931 GCTCTGGCCTTCGGTATGGGACCACTTGGTCTCTTCAGGGAA-----TATTCTATT 984  
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
 Db 985 GGACAAGTACTCACTGTATTCTTTCTGTATTAAATTGGGGCTTTAGTGTGGACAGGCA 1044  
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
 Db 1045 TCTCCAAGCATTGAAGCATTTCCAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 1104  
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
 Db 1105 ATTGATAAATAAGCCAAGTATTGACAGCTTTCGAAGAGTGGGCACCAACAGATAATT 1164  
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
 Db 1165 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAGAAGTTAAG 1224  
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
 Db 1225 ATCTGAAGGGCTGAACCTGAAGGTGGCAGAGTGGCGAGCGTGGCCCTGGTTGGAAC 1284  
 QY 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
 Db 1285 AGTGGCTGTGGGAAGAGCACACAGCTCCAGCTGATGAGAGGCTCTATGACCCACAG 1344  
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
 Db 1345 GGGATGGTCAGTGTGTGTGGACAGATATTAGACCATAATGTAGGTTTCTACGGGAA 1404  
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
 Db 1405 ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTGCCACCACGATAGCTGAAACATT 1464  
 QY 284 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 303  
 Db 1465 CGTATGGCCGTGAAATGTCACCATGTGATGAGATGATGAGAAAGCTGTCAAGGAAGCCA 1524  
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
 Db 1525 GCCTATGACTTTATCATGAAACTGCTCATTAATTTGACACCTCTGTTGGAGAGAGGG 1584  
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
 Db 1585 GCCCAGTTGAGTGGTGGCAGAGAGAGGATGCCCATTCACGTGCGCTTGGTTCCGCAAC 1644  
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
 Db 1645 CCCAAGATCCTCTGCTGGATGAGGCCACGCTGAGCTTGGACACAGAAACGGAAGCAGT 1704  
 QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
 Db 1705 GTTCAGTGGCTCTGGATAAGGCCAGAAAGGTGGACCAACCATTTGTAGTATGCTCATCGT 1764  
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403

Db 1765 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATTTGGAG 1824  
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer 423  
Db 1825 AAGGAAATCATGATGAATCATGAAAGAAAGGCATTTACTTCAAACTGTGCAAAATG 1884  
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
Db 1885 CAGACAGCAGGAAATGAAGTTGAATAGAAATGCAAGTCAATCCAAAGTGAAT 1944  
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 1945 GATGCTTGGAAATGCTTCAAAATGATTCAAGATCCAGTCTTAATAAGAAAGATCAACT 2004  
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 470  
Db 2005 CGTAGGAGTGTGCTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
Qy 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 490  
Db 2065 GATCAAAAGTATACCTCCAGTTCTCTTTTGAGGATATGAAGCTAAATTTAACTGAATGG 2124  
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2125 CCTATTCTTGTGTGTATTTTGTGCTATTAATAAGAGCCCTGCACCAAGCAATTT 2184  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2185 GCATAATAATTTCAAGATATAGGGTTTACAGAAGATTGATGATCTGAAACAAAA 2244  
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2245 CGACAGAATAGTAATCTGTTTCTCACTATTGTTCTAGCCCTTGGAAATATTTCTTTATT 2304  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2305 ACATTTTCTTCAAGGTTTTCACATTTGGCAAAAGCTGGAGAGATCTCCACCAAGCGGCTC 2364  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleTyrPheAspGluLysGlu 589  
Db 2365 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCTAAA 2424  
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2425 AACACCACTCGAGCATTTGACTACCGAGCTCGCCAATGATGCTCAAGTTAAAGGGGCT 2484  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2485 ATAGGTTCCAGGCTGTGTTAAATTTACCCAGAATATAGCAAAATCTTGGGACAGGAATTAAT 2544  
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2545 ATATCTTCAATCATGTTGGCAACTAACACTGTTACTCTTAGCAATGTGACCCATCATTT 2604  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2605 GCAATAGCAGGAGTTGTTGAAATGTGTCTCGCAAGCACTGAAAGATAAGAAA 2664  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2665 GAAC TAGAAGGTGTGGGAAAGATCGCTACTGAAAGCAATAGAAAACCTTCCGAACCGTTGTT 2724  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 2725 TCATTGATCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 2784  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 2785 AGAACTCTTTGAGAAAGCACACATCTTTGGAATTACATTTTCTTCCACCAGCAATG 2844  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749  
Db 2845 ATGTATTTTCTTCTATGCTGATGTTTCCGGTTTGGAGCTACTTGGTGCCACATAAACTC 2904

Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 2905 ATGAGCTTTGAGGATGTTCTGTAGTATTTCAGCTGTTGTCTTTGGTGCATGGCCGTG 2964  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysLysSerGlyAlaAlaHisLeu 789  
Db 2965 GGGCAAGTCAGTTCAATTTGCTCTGCTATGCAAGCCAAAGCCAAATATATCAGCAGCCACATC 3024  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3025 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGTACAGCAGGAGGCCCTAATGCCG 3084  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 3085 AACACATTGGAAGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3144  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3145 GACATCCCAGTGTCTTCAGGACTGAGCCTGGAGGTGAAGAGGGCCAGACGCTGGCTCTG 3204  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3205 GTGGGAGCAGTGGCTGTGGAGAGACACAGTGTCTCCAGCTCTCTGGAGCGGTTCTACGAC 3264  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAlaLysGlyLeuLeuAsnValGlnTyr 889  
Db 3265 CCCTTGGCAGGAAAGTGTCTGTGTCGCAAGAAATAAAGCGACTGAATGTTCACTGGT 3324  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3325 CTCGAGACACCTGGGCATCGTGTCCCAGGAGCCCATCTCTGTTTCACTGCACATGTCT 3384  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929  
Db 3385 GAGAACATTGCTTATGAGACAAACAGCCGGTGTGTACAGGAGAGATCTGTGAGGGCA 3444  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3445 GCAAGGAGGCCAACATACATACGCTTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3505 GTAGGAGACAAAGGAACTCAGCTCTCTGTGGTGGCCAGAAAACACGCAATTGCCATGCTCT 3564  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 3565 GCCCTTGTAGCAGCCTCATATTTTGTGATGAAGCCACGCTCAGCTCTGGATACA 3624  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 3625 GAAAGTGAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGACAGAGCGCCACCTGCATT 3684  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 3685 GTGATTGCTCACCGCTCTCCACCACTCCAGAAATGCAAGCTTAATAGTGGTGTTCAGAAT 3744  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGCACTATTTT 3804  
Qy 1050 LysLeuValAlaGlnSer 1056  
Db 3805 TCAATGTCAGTGTCCAGGCT 3825  
RESULT 10  
AAZ49333  
ID AAZ49333 standard; cDNA; 3860 BP.  
XX  
AC AAZ49333;  
XX  
DT 14-MAR-2000 (first entry)  
XX

Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.

Multi:drug resistance; MDR-1; P-glycoprotein;  
transmembrane efflux pump; haematopoietic stem cell; transduction;  
bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
gene therapy; gene replacement; genetic defect; thalassaemia;  
Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
cytokine; mutant; ds.

Synthetic.

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Key Location/Qualifiers  
CD5 1.3843  
/\*tag= a  
/product= "Human G185V mutant MDR-1 protein"

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Ex vivo expansion of hematopoietic stem cells transduced with a  
sequence encoding human multidrug resistance-1, used for bone marrow  
transplantation -  
Example 1; Page 79-82; 113pp; English.  
This sequence represents cDNA encoding human G185V mutant multidrug  
resistance protein MDR-1, where the Gly residue at position 185  
of the wild-type protein (AAV58186) is replaced by Val. MDR-1 is a  
transmembrane efflux pump, responsible for the export of drugs from  
cells, particularly cancer cells. The wild-type MDR-1 shows increased  
resistance to etoposide and decreased resistance to vinca alkaloids  
compared with the G185V mutant. The invention relates to transducing  
haematopoietic stem cells with nucleic acid encoding an MDR protein  
and culturing the modified cells. The modified haematopoietic stem  
cells are useful in bone marrow transplantation (to reconstitute  
haematopoietic systems in patients who have undergone chemotherapy or  
radiation therapy) and in ex vivo gene therapy of genetic defects in  
cells derived from haematopoietic stem cells, e.g., thalassaemia,  
Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
cells can also be used to identify factors involved in regulating  
proliferation and differentiation in haematopoietic stem cells.  
Haematopoietic stem cells that express MDR-1 will be protected against  
chemotherapeutic agents, so can be engrafted while the patient is  
undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
provides sufficient cells to permit standard biochemical analysis.  
Overexpression of MDR-1 allows cytokine-driven expansion of  
haematopoietic stem cells by at least 10-fold compared with a maximum  
of 4-fold in known procedures.

Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:

Pred. No.:	1.88e-282	Length:	3860
Score:	3100.50	Matches:	595
Percent Similarity:	75.54%	Conservative:	211
Best Local Similarity:	55.76%	Mismatches:	218
Query Match:	58.17%	Indels:	43
DB:	21	Gaps:	6

US-09-873-409-4 (1-1058) x AAZ49333 (1-3860)

QY	4	SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluGluVal	23
Db	712	TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC	771
QY	24	LeuSerSerIleAArgThrValleAlaPheAArgAlaGlnGlnLysGluLeuGlnAArgSer	43
Db	772	TTGGCAGCAATTAGAAGCTGCTGATTTGCAATTTGGAGGCAAAAGAAAGAACTTGAA	825
QY	44	PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys	63
Db	825	-----	825
QY	64	ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	83
Db	826	-----AGGTACACAAATAATTTAGAAGAAGCTAAAGAAATTTGGATAAAG	870
QY	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	103
Db	871	AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGCTAT	930
QY	104	GlyLeuAlaPheThrTyrGlyThrSerIleLeuAsnGlyGluProGlyTyrThrIle	123
Db	931	GCTCTGGCCTTCGTGATGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT	984
QY	124	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	143
Db	985	GGACAAGTACTCACTGTAATCTTTCTGTATTAAATGGGGCTTTTAGTGTGGACAGGCA	1044
QY	144	ValProHisPheGluThrPheAlaIleAlaAArgGlyAlaAlaPheHisIlePheGlnVal	153
Db	1045	TTCTCCAAGCATTTGCAATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGTA	1104
QY	164	IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	183
Db	1105	ATTGATAAATGAAGCCAGTATTGACAGCTATTGCAAGAGTGGGCACAAACAGATAATTT	1164
QY	184	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
Db	1165	AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTCACCATCTCGAAAGAGTTAAG	1224
QY	204	IleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	1225	ATCTGAAGGGCTGAACCTGAAGGTGCAGTGGCAGCGTGGCCCTGGTGGTGAAC	1284
QY	224	AsnGlySerGlyLysSerThrValValGlnLeuGlnAArgLeuTyrAspProAspAsp	243
Db	1285	AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGACAGAGGCTCTATGACCCACAGAG	1344
QY	244	GlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnValArgHisTyrArgAsp	263
Db	1345	GGATGGTCACTGTTGATGGACAGGATATTAGGCATTAATGTAAGGTTTCTACGGAA	1404
QY	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	283
Db	1405	ATCATTTGGTGTGAGTCAGGAACCTGTATTGTTGCCACCATGATGATGATGATGATGAT	1464
QY	284	LysTyrGlyArgAspValThrAspGluGluMetGluAArgAlaAArgGluAlaAsn	303
Db	1465	CGTATGGCGGTGAAATGTCCACCATGATGATGATGATGATGATGATGATGATGATGAT	1524
QY	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323
Db	1525	GCCTATGACTTTTATCATGAACCTGCCTCATTAATTTGACACCCCTGGTGGAGAGAGGG	1584
QY	324	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaAlaValAlaValArgAsn	343
Db	1585	GCCAGTTGAGTGTGGGCAAGAGCAGAGGATGCCATTCAGCTGGCCCTGGTGGTGGCAAC	1644
QY	344	ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1645	CCCAAGATCTCTCTGCTGGATGAGGCCACCTGAGCTTGGACACAGAAAGGAGCAGTG	1704

QY 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 393  
Db 1705 GTTCAGGTGCTCTGGATAGGCAGAAAAGGTCGGACCACTTGTGATAGCTATCGT 1764  
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1765 TTGTCTACAGTTCGTAACTGCTGAGCTCATCGCTGTTTCGATGATGAGTCATGTGGAG 1824  
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 1825 AAAGGAATCATGATGAACCTCATGAAGAGAAAGGCACTTACTTCAAACTTGTCAACAATG 1884  
QY 424 Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr 437  
Db 1885 CACAGCAGCAAGAAATGAAGTTGAATAGAAATGACAGCTGATGAATCCAAAGTGAAT 1944  
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 1945 GATGCCTTGGAAATGCTTCAATGATTCAAGATCCAGTCTAAATAGAAAAGATCAACT 2004  
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470  
Db 2005 CGTAGGAGTCTCGGTGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
QY 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490  
Db 2065 GATGAAGATATACCTCCAGTTCCTTTTGAGGATTATGAAGCTAAATTTAACTGAATGG 2124  
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2125 CCTATTGTTGTTGGTGTATTGTTGGCTATTATAATGGAGCCCTGCAACAGCAATTT 2184  
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2185 GCAATRAATTTTCAAAGATTATAGGGTTTTACAAAGAAATTTGATGATCCTGAAACAAA 2244  
QY 530 LysHisAspAlaGluIleLysThrMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2245 CGACAGAATAGTAACCTGTTTTCACATTTGCTTACCTGCTAGCCCTTGGAAATTTCTTTAT 2304  
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2305 ACATTTTCTTCAAGTTTCACATTTGGCAAGCTGGAGAGATCTTCAACCAAGCGGCTC 2364  
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGlyLysGlu 589  
Db 2365 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCTAAA 2424  
QY 590 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2425 AACACCACTGGAGCATTTGACTACAGGCTCGCAATGATGCTCAAGTTAAAGGGCT 2484  
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2485 ATAGTTCACAGGCTTGTGTAATTTACCAGAAATATAGCAATCTTGGGACAGGAATAATT 2544  
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValIle 649  
Db 2545 ATATCTTCACTATGTTGGCAACTAACACTGTTACTCTTAGCAATGTACCAATCAT 2604  
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2605 GCAATAGCAGGAGTTGTTGAAATGMAAATGTTGCTGGCAAGCACTGAAAGATAGAAA 2664  
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2665 GAACTAGAAGGTGCTGGGAAGATGCTGACTGAAGCAATAGAAAATCTCCGAACCGTTGTT 2724  
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuThrGlnHis 709  
Db 2725 TCTTTGACTCAGAGCAGAGATTGAAATATGATGCTCAGAGTTTGCAGGTACCAATAC 2784  
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729

Db 2785 AGAAACTCTTTTGGAAAGCACACATCTTTTGGAAATTTACATTTTCTTCACCCAGGCAATG 2844  
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749  
Db 2845 ATGTAATTTTCTATGCTGATGATGTTCCGGTTTGGAGCTACTTGGTGGCACAATAAATC 2904  
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTCAGCTGTTGCTTGTGTCATGGCCGTG 2964  
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 2965 GGCACAGTCAGTTCATTTGCTCTGACTATGCAAGCCAAATATATCAGCAGCCACATC 3024  
QY 790 PheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3025 ATCATGATCATTTGAAAAACCCCTTTGATTCAGCTCAGCAGCAGGAGCCCTAATGCCG 3084  
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 829  
Db 3085 AACACATTGGAAGAAATGTCATTTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3144  
QY 830 AspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3145 GACATCCCATGTTCTCAGGAGCTGAGCTGGAGGTGAAGAGGCCAGCGCTGGCTCTG 3204  
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3205 GTGGGACAGAGTGGCTGTGGGAAGACACAGTGTCTGAGCTCTGAGCGGTCTACGAC 3264  
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAlaAlaLysGluLeuAsnValGlnTyr 889  
Db 3265 CCCTTGGCAGGAAAGTGTCTGTATGTGCAAAAGAAATAAAGCGACTGAATGTTAGTGG 3324  
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3325 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGCACATGCT 3384  
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3385 GAGAAATTCCTATGAGACACACAGCCGGTGTGTCTACAGGAAGAGATCTGTAGGGCA 3444  
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3445 GCAAGGAGGCCCAACATACATGCTTTCATCGACTCTGCTTAATAATATAGCACTAAA 3504  
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3505 GTAGGAGACAAAGAACTCAGCTCTCTGTTGGCCAGAAACACAGCATTTGCCATAGCTCG 3564  
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 3565 GCCCTTGTAGACGCTCATATTTTGTGTTGGATGAAGCCACGTCAGCTCTGGATACA 3624  
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 3625 GAAAGTGAAAGGTTCTCCAAGAGCCCTGGCAAGCCAGAGAGCCGACCTGCATT 3684  
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 3685 GTGATTGCTCACCGCTCTCCACCATCCAGATTCAGACTTAATAGTGTGTTTTCAGAAAT 3744  
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGAGCTGCTGGCACAGAAAGCATCTATTTT 3804  
QY 1050 LysLeuValAsnAlaGlnSer 1056  
Db 3805 TCATGCTCAGTCTCCAGCT 3825  
RESULT 11  
ABA94365

ID ABA94365 standard; DNA; 3860 BP.

XX AC ABA94365;

XX DT 26-MAR-2002 (first entry)

XX Human BCRP DNA related seq id No. 1.

DE DE Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;  
KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;  
KW cardiant; gene therapy; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1..3843

FT /\*tag= a

XX WO200192877-A2.

PN 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17459.

XX 31-MAY-2000; 2000US-0584586.

PR 29-MAY-2001; 2001US-0866866.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Sorrentino B, Schuetz J;

XX WPI; 2002-114368/15.

DR P-PSDB; ABB07266.

XX Identifying a stem cell, for treating e.g., muscular dystrophy,  
PT myocardial infarction, Parkinson's disease, or neurodegenerative  
PT disorders, comprises detecting the expression of an ATP transport  
PT protein (BCRP) by a cell -

PS Disclosure; Page 53-55; 87pp; English.

XX The invention provides a method of identifying and/or isolating a stem  
CC cell that involves detecting the expression of an ATP transport protein  
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample  
CC comprising stem cells. The isolated stem cells may be used in the  
CC treatment of diseases such as muscular dystrophy, degenerative liver  
CC disorder, myocardial infarction, Parkinson's disease, degenerative  
CC disorders of the brain, and for tissue regeneration or replacement.  
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for  
CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
CC also be used as cell targets in gene therapy protocols. The present  
CC sequence represents a sequence related to the BCRP for which no relevant  
CC information has been provided in the specification.

SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:

Pred. No.: 1.88e-282 Length: 3860  
Score: 3100.50 Matches: 595  
Percent Similarity: 75.54% Conservative: 211  
Best Local Similarity: 55.76% Mismatches: 218  
Query Match: 58.17% Indels: 43  
DB: 24 Gaps: 6

US-09-873-409-4 (1-1058) x ABA94365 (1-3860)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23

Db 712 TCATTACTGATAAAGAACTCTTAGCGTATGCCAAGCTGAGCAGTAGCTGAAGAGTC 771

Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db 772 TTGGCAGCAATTAGAACTGTGATTGCAATTTGGAGGACAAAAGAAAGAACTTGA----- 825  
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
Db 825 ----- 825  
Qy 64 ValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 826 -----AGGTACAAACAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870  
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
Db 871 AAGAGCTATTACAGCCCAATATTTCTATAGTGTCTCTTCTCTGCTCTATCATCTTAT 930  
Qy 104 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
Db 931 GCTCTGGCCTTCGGTATGGACCACCTTGGTCTCTCAGGGAA-----TATTCTATT 984  
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAala 143  
Db 985 GGACAAAGTACTCACTGTATTCTTTCTGATTAAATTTGGGGCTTTTAGTGTGGACAGCA 1044  
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 1045 TCTCCAAGCATTGAAGCATTTCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104  
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 1105 ATTGATAAATGAAGCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACAGATAATA 1164  
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
Db 1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAAGAGTTAAG 1224  
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1225 ATCTTGAAGGGCTGAACCTGAAGTGAAGTGGGACAGCGTGGCCCTGGTGGAAAC 1284  
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243  
Db 1285 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGACAGGCTCTATGACCACAGAG 1344  
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1345 GGGATGTCAGTCTGATGGACAGGATATTAGGACCAATAATGTAAGTTCCTACGGGAA 1404  
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1405 ATCATTTGGTGTGGTGAAGTCAAGAACCTGTATTGTTGCCACCAAGTAGCTGAAAACATT 1464  
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1465 CGCTATGGCGTGAATATGTCACCATGATGAGATGAGAAAGCTGTCAAGGAAGCCCAT 1524  
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
Db 1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 1584  
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1585 GCCCAGTTGAGTGTGGCAGAAAGCAGAGGATGCCCATTCAGCTGCCCTGGTTCCGAAC 1644  
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1645 CCCAAGATCCTCTGCTGGATGAGGCCACCTCAGCCTTGGACACAGAAAGCAGCAGTG 1704  
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 1705 GTTCAGTGGCTCTGGTAAGGCCAGAAAAGTTCGACCAACCATTTGTGATAGTCTCATCGT 1764  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1765 TTGCTACAGTTTCGTAATGCTGACGTCATCGCTGCTGTTTCGATGATGAGTATTGTGGAG 1824



QY	404	LyGlyAlAHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	423
DB	1825	AAAGGAAATCATGATGAATCTATGAAGAGAAAGCAATTTACTTCAAACCTGTGCACAAATG	1884
QY	424	Gln-----AspilelyslsAlaAspGluGlnMetGluSerMetThr	437
DB	1885	CAGACAGCAGGAAATGAAGTTGAATTTAGAAAATGCAGCTGGATGAATCCAAAAGTGAAT	1944
QY	438	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	452
DB	1945	GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGAAAAGATCAACT	2004
QY	453	---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys	470
DB	2005	CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAAGAGCGCTCTG	2064
QY	471	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTirp	490
DB	2065	GATGAAGATATACCTCCAGTTCCTTTTGGAGGATATGAAGCTAAATTTAACTGAATGG	2124
QY	491	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	510
DB	2125	CCTATTATTGTTGGTGTATTTTGTGCCATATAAATGAGAGCGCTGCACACAGCAATTT	2184
QY	511	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	529
DB	2185	GCAATAATATTTCAAAGATATATAGGGTTTTTACAAGAAATGATGATCCTGAAACAAAA	2244
QY	530	LySHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	549
DB	2245	CGACAGAATAGTAACCTGTGTTTCACTATTGTTCTCAGCCCTGGAAATATTATTCTTTTATT	2304
QY	550	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	569
DB	2305	ACATTTTTCTCCAGGTTTCATTTTGCAAGCTGGAGAGATCCTCACCACAGCGGCTC	2364
QY	570	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirpPheAspGluLysGlu	589
DB	2365	CGATACATGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCATAA	2424
QY	590	AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	609
DB	2425	AACACACCTGGAGCATTGTACTACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCT	2484
QY	610	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	629
DB	2485	ATAGTTTCCAGGCTTCTGTATATCCCAAGATATAGCAAAATCTTGGGACAGGAATTAAT	2544
QY	630	IleSerPheIleTyrGlyTirpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	649
DB	2545	ATATCCTTCATCTATGTTGGCAACTAACACCTGTGTACTCTTAGCAATGTGTACCCATCAT	2604
QY	650	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	669
DB	2605	GCAATAGCAGGAGTGTGAAATGAAAATGTTGTCTGGCAAGCACTGAAAAGATAAGAAA	2664
QY	670	GluLeuLysHisIalaglyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
DB	2665	GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCCGAACCGTTGTT	2724
QY	690	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	709
DB	2725	TCITTTGACTCAGAGCAGGAAGTTTGACATATGATGCTCAGAGTTTCAGGTACCATATAC	2784
QY	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	729
DB	2785	AGAAACTCTTTTGAGGAAGACACACATCTTTTGGAAATTTACATTTTCTCTCACCCAGGCAATG	2844
QY	730	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	749
DB	2845	ATGTATTTTTTCCATGCTGGATGTTTCCGGTTTGGAGGCTACTTGTGGGCACATAAACTC	2904

Qy	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	769
Db	2905	ATGAGCTTTTGAGATGTCCTGGTATTATTTTCAGCTGTCTTCTGGTGCCATGCCGCGT	2964
Qy	770	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	789
Db	2965	GGCAGACTCAGTTCAATTTGCTCTGACTATGCCAAGGCATAATATCATGACGCCCAATC	3024
Qy	790	PheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGlnGlyLysLeuPro	809
Db	3025	ATCATGATCATTCGAAACCCCTTTGATTGACACTACAGCACGGAAGGCTTAATGCCG	3084
Qy	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829
Db	3085	AACACATTTGGAAGGAATGTCATTTGGTGAAGTTGATTCAACTATCCCACCCCGACCG	3144
Qy	830	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	849
Db	3145	GACATCCAGTAGCTTCAGGACTGAGCTGGAGGTGAAGAAGGCCAGACCTGGCTCTG	3204
Qy	850	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
Db	3205	GTGGCGCAGCTGGCTGTGGGAAGCACAGTGGTCCAGCTCTCTGGAGCGGTCTTACGAC	3264
Qy	870	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	889
Db	3265	CCCTTGGCAGGAAAGTGTGCTGTATGTCGAAAGAAAATAAAGCGCACTGAATGTTCACTGG	3324
Qy	890	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	909
Db	3325	CTCCGAGCACACTCGGCATCGTGTCCAGGAGCCCCCTCTGTTTGACTGCAGCATTTGCT	3384
Qy	910	GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla	929
Db	3385	GAGAACATTGCTTATCGAGACACACCGCGGTGTGTACAGGAAGAGATCGTCAGGGCCA	3444
Qy	930	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	949
Db	3445	GCAAAGGAGGCCAACATACATGCTTCATCGAGTCACTGCCTTAATAAATATAGCACTAAA	3504
Qy	950	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	969
Db	3505	GTAGGAGACAAGAAACTCAGCTCTCTGTGTGCCAGAAACACCGCANITGGCCATAGCTCGT	3564
Qy	970	AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	989
Db	3565	GCCCTGTTAGACAGCTCATATTTTGGTATGAAGCCACGTCAGCTCTGGATACA	3624
Qy	990	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1009
Db	3625	GAAAGTGAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGGCCGACACCTGCATT	3684
Qy	1010	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn	1029
Db	3685	GTGATTGCTCACCGCTGTCCACCATCCAGATCGAGACTTAATAGTGGTGTTCACGAAT	3744
Qy	1030	GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuLeuLeuLeuLeuLeuLeuLeu	1049
Db	3745	GGCAGATCAAGAGCATGGCAGCATACAGCTGTCTGGCACAGAAAGGCATCTATTTTT	3804
Qy	1050	LysLeuValAsnAlaGlnSer 1056	
Db	3805	TCAATGTCAGTTCAGGCT 3825	
	RESULT 12		
	ABA94366		
ID	ABA94366 standard; DNA; 3860 BP.		
XX	ABA94366;		
XX			
DT	26-MAR-2002 (first entry)		
XX			
DE	Human BCRP DNA related seq Id No. 3.		

DT 26-MAR-2002 (first entry)

Human BCRP DNA related seq Id No. 3.



XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;  
 KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;  
 KW cardiant; gene therapy; ds.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH 1..3843  
 CDS /\*tag= a

XX WO200192877-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17459.

XX 31-MAY-2000; 2000US-0584586.

XX 29-MAY-2001; 2001US-086866.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Sorrentino B, Schuetz J;

XX WPI; 2002-114368/15.

XX P-PSDB; ABB07267.

XX Identifying a stem cell, for treating e.g., muscular dystrophy,  
 PT myocardial infarction, Parkinson's disease, or neurodegenerative  
 PT disorders, comprises detecting the expression of an ATP transport  
 PT protein (BCRP) by a cell

XX Disclosure; Page 59-60; 87pp; English.

XX The invention provides a method of identifying and/or isolating a stem  
 CC cell that involves detecting the expression of an ATP transport protein  
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample  
 CC comprising stem cells. The isolated stem cells may be used in the  
 CC treatment of diseases such as muscular dystrophy, degenerative liver  
 CC disorder, myocardial infarction, Parkinson's disease, degenerative  
 CC disorders of the brain, and for tissue regeneration or replacement.  
 CC Haematopoietic cells can be used in bone marrow transplants (e.g., for  
 CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
 CC also be used as cell targets in gene therapy protocols. The present  
 CC sequence represents a sequence related to the BCRP for which no relevant  
 CC information has been provided in the specification.

XX SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:

Pred. No.: 1.88e-282 Length: 3860  
 Score: 3100.50 Matches: 595  
 Percent Similarity: 75.54% Conservative: 211  
 Best Local Similarity: 55.76% Mismatches: 218  
 Query Match: 58.17% Indels: 43  
 DB: 24 Gaps: 6

US-09-873-409-4 (1-1058) x ABA94366 (1-3860)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23

Db 712 TCATTTACTGATAAGAACTCTTAGCGTATGCAAAAGCTGGACAGTAGCTGAAGAGTC 771

QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgSer 43

Db 772 TTGCGACCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA----- 825

QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db 825 ----- 825

QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83

Db 826 -----AGGTACAAACAAAAATTTAGAAAGAGCTAAAAAGAAATTCGGATAAAG 870  
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
 Db 871 AAAGCTATTACAGCCAAATTTCTATAGGTGCTGCTTCTGCTGATCTATGCATCTTAT 930  
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123  
 Db 931 GCTCTGGCCTCTCTGGTATGGGACCACCTTGGTCTCTCAGGGGAA-----TATTCTATT 984  
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
 Db 985 GGACAGTACTCACTGTATTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1044  
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
 Db 1045 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1104  
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
 Db 1105 ATTGATAATAAGCCAGTATTGCAGCTATTGCAAGAGTGGGCACAAACACAGATAATT 1164  
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
 Db 1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1224  
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
 Db 1225 ATCTTGAAGGGCTGAACCTGAAGTGCAGAGTGGGCACAGCGTGGCCCTGGTGGAAAC 1284  
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243  
 Db 1285 AGTGGCTGTGGGAGAGCACACAGTCCAGCTGATGCGAGAGGCTCTATGCCCCACAGAG 1344  
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
 Db 1345 GCGATGCTCAGTGTTCATGGACAGAGATATTAGCACCAATAATGTAAGGTTTCTACGGGA 1404  
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
 Db 1405 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1464  
 QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
 Db 1465 CGCTATGGCGGTGAAATGTCCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAAT 1524  
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
 Db 1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGCACACCTGTTGGAGAGAGAGGG 1584  
 QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
 Db 1585 GCCAGTTGAGTGGTGGGAGAGAGAGAGATCGCCATTCGACGTGCCCTGGTGGTGGCAAC 1644  
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
 Db 1645 CCCAGATCTCTCTGCTGGATGGGCCACGTCAGCTTGGACACAGAAAGCGAGCAGTG 1704  
 QY 364 ValGlnAlaAlaLeuGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
 Db 1705 GTTCAGTGGCTCTGGATAAGGCCAGAAAAGGTCCGACACCATTTGTGATGCTCATCGT 1764  
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
 Db 1765 TTCTCTACAGTTCGTAATGCTGACGTGCTGCTGGTTCGATGATGAGTCATTTGGAG 1824  
 QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
 Db 1825 AAAGGAATCATGATGAACCTCAAGAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 1884  
 QY 424 Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr 437

Db 1885 CACACAGCAGGAATGAAGTTGAATTAGAAAATGACGCTGATGAATCCAAAAGTGAAT 1944  
Qy 438 TysSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 1945 GATGCCTTGGAAATGCTCTCAAAATGATTCAGATCCAGTCTAATAAGAAAAAGCAACT 2004  
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470  
Db 2005 CGTAGAGAGTGCCTGGATCACAAGGCCAAGACAGAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
Qy 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTTP 490  
Db 2065 GATGAAGATATACCTCCAGTTTCTTTGGAGGATATGAAGCTAAATTTAACTGAATGG 2124  
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2125 CCATTATTTTGTGTGTGTTTGTGCTATTAATGAGGCTGCAACAGCATTT 2184  
Qy 511 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2185 GCAATAATATTTTCAAAGAGATTATAGGGGTTTTTACAAGAAATTGATGATCTGGAACAAA 2244  
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2245 CGACAGAAATAGTAACCTGTTTCTACTATGTTTCTAGCCCTTGGAAATATTTCTTTTAT 2304  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2305 ACATTTTCTTCAAGGTTTCATATTGGCAAGCTGGAGAGATCTCCACCAAGCGCTC 2364  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589  
Db 2365 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGTATGACCCCTAAA 2424  
Qy 590 AsnSerThrClyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2425 AACACCACTGGAGCATGTAGTACAGGCTCGCAATGATGCTCTCAAGTTAAAGGGCT 2484  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2485 ATAGTTTCCAGGCTGCTGTATATACCCAGAAATATAGCAATCTTGGACAGAAATAT 2544  
Qy 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2545 ATATCTTCTATCTATGTTGGCAACTAACACTGTACTCTTAGCAATTTGACCCATCAT 2604  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2605 GCAATAGCAGGAGTTTGAATGAAATGTTTGTCTGGACAAGCACTGAAAGATNAAGAA 2664  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2665 GAATAGAGGTGCTGGGAAGATCGTACTGAAGCAATAGAAAACCTTCCGAACCGTTGT 2724  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 2725 TCTTTGACTCAGGAGCAGAAATTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 2785 AGAAATCTTTGAGGAAGACACATCTTTGGAAATATACATTTTCTTCCACCCAGGCAATG 2844  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749  
Db 2845 ATGTATTTTCTTATGCTGATGTTCCGGTTTGGAGCTTACTTGTGGCACATAAATC 2904  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 2905 ATGAGCTTTCAGGATGTTCTGTTAGTATTTTACGTCGTTTGGTGGCCGCGTG 2964  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 2965 GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAGCCAAAATATATCAGCAGCCCATC 3024

Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlnGlyLysLysPro 809  
Db 3025 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAAGCCCTAATGCCG 3084  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 3085 AACACATTGGAAGAAATGTCACTTTGTGTGAAGTTGTATTCAACTATCCACCCGACCG 3144  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3145 GACATCCAGTCTTCCAGGACTGAGCTTGGAGTGAAGAGGCCAGACGCTGGCTCTG 3204  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3205 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCGAGGGTGTCTACGAC 3264  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTTP 889  
Db 3265 CCCTTGGCAGGAAAGTGTCTTGTATGCAAAAGAAATAAAGCGACTGAATGTTCACTGG 3324  
Qy 890 LeuAsnSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3325 CTCGAGACACCTGGGCATCGTGTCCAGAGCCCATCTCTGTTGACTGCAGCATTGCT 3384  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3385 GAGAACATTTGCTATGGAGACAACACGCGGTGTGTACAGGAAGAGATCGTGAGGGCA 3444  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3445 GCAAGAGGAGCCCAACATACATCGCTTTCATCGAGTCACTGCTTAATAAATATAGCATAAA 3504  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3505 GTAGGAGACAAGAACTCAGCTCTCTGTGGCCAAACACAGCAATGGCATAGTCTGT 3564  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 3565 GCCTTGTTAGACAGCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACA 3624  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 3625 GAAAGTGAAGAGTTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3684  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 3685 GTGATTGCTCACCGCTGTCCACCATCCAGATGCAAGACTTAATAGTGGTGTTCAGAA 3744  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTCTGGCACAGAAAGCATCTATTTT 3804  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 3805 TCAATGGTCAGTGTCCAGGCT 3825  
RESULT 13  
AAF86127 standard; cDNA; 4186 BP.  
ID AAF86127  
XX AAF86127;  
AC AAF86127;  
XX 25-JUN-2001 (first entry)  
DT  
XX Cynomologous monkey P-glycoprotein cDNA.  
DE Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL;  
XX Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL;  
KW efflux pump; ss.  
XX Macaca fascicularis.  
OS  
XX Key Location/Qualifiers  
FH

```
FT CDS 100..3942
FT /*tag= a
FT /product= "pgp"
FT /note= "p-glycoprotein"
```

WO200123565-A1.

05-APR-2001.

28-SEP-2000; 2000WO-US26592.

28-SEP-1999; 99US-0156921.

12-OCT-1999; 99US-0158818.

(GENT-) GENTEST CORP.

Stocker PJ, Steimel-Crespi DT, Crespi CL;

WPI: 2001-316136/33.

P-PSDB; AAB81064.

Novel isolated nucleic acid encoding cynomolgous monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting bioavailability of compound and increasing PGP transporter activity in cell -

Example 1; Page 51-57; 84pp; English.

This invention relates to a polynucleotide sequence encoding a cynomolgous monkey P-glycoprotein (PGP), and an allelic variant of the PGP protein. PGP, also known as multidrug transporter, MDR1 is a member of the ABC transporter superfamily. The enzyme serves as an efflux pump of exporting small molecules across the cell membrane. The invention includes a cynomolgous monkey (Macaca fascicularis) PGP coding sequence and protein, and also that of an allelic variant. The PGP polynucleotide sequence is useful for increasing PGP transporter activity in a cell. Antisense sequences of the cDNA are useful for inhibiting PGP transport activity in a mammalian cell. They may also be used for increasing the bioavailability of a drug. The present sequence represents cDNA encoding cynomolgous monkey P-glycoprotein.

Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:		
Pred. No.:	2,125-282	Length: 4186
Score:	3100.50	Matches: 594
Percent Similarity:	75.4%	Conservative: 216
Best Local Similarity:	55.5%	Mismatches: 216
Query Match:	58.1%	Indels: 47
DB:	22	Gaps: 7

US-09-873-409-4 (1-1058) x AAF86127 (1-4186)

4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23

Db 811 TCATTCTGATAAGAACTCTTAGCTTATGCCAAAGCTGGAGCAGTAGCTGAAGAGGTC 870

Ov 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db 871 TTGGCAGCAA TTAGAACTGTGATTGGAGGACAAAGAAAGAA CTCGAA ---- 924

Qv 44 pheLeuLeuAsnIleThrArgTyrAlaTmPheTyrPheProGlnTmPheLeuSerCys 63

nb 924 ----- 924

64 ValLeu\*\*\*pHeValArgTyrThrGlnAsnLeuIleValAspAspAlaValAspPheGlyTleVal 83

[illegible]

84 **NR** **NR** Thr I Leu Leu Ser Leu Cys Val Ala Val Tyr Phe Phe Met Asn Gln Val Thr Tyr 103

[illegible]

```
Db 2104 CGTAGGAGTCCGTCGGATCACAGGCCAA-----GACAGAAAGCTTAGTACCAAGAG 2157
QY SerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db 2158 GCTCTGGATGAAAGTATACCTCCAGTTTCCTTTGGAGGATTATGAAGCTAAATTAAC 2217
QY GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db 2218 GAGTGGCCTTATTTGTGTGTGTATTTGTGCCATTATAAATGAGGCTCGCAACCA 2277
QY ValPheSerIleIlePheAlaLysIleIleThrMetPhe-----GlyAsnAsnAspLysThr 527
Db 2278 GCATTTCAGTAAATATTTCAAGATTATAGGATTTTTCACAAATGATGATGCCGAA 2337
QY ThrLeuLysHisAspAlaGluIleIleSerMetIlePheValIleLeuGlyValIleCys 547
Db 2338 ACAAAACGACAGATAGTACTGTTTCTACTATTGTTCTAGTCTTGAATGTGTTCT 2397
QY PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db 2398 TTTATTACATTTTCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAG 2457
QY ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db 2458 CGCTCCGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGCTGTTTGTATGAC 2517
QY LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db 2518 CCTAAAACACCACTGGAGCATTTGACTACAGGCTGCCAATGATGCTGCTCAAGTTAA 2577
QY GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db 2578 GGGGCTATAGTTTCCAGGCTTGCTATAATTTACCCAGAAATATAGCAAAATCTTGGACAGGA 2637
QY ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db 2638 ATAAATTATATCTTAATCTATGTGTGGCACTGACACTGTACTCTTAGCAATGTATGCC 2697
QY ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db 2698 ATCATTTGCAATACAGGAGTTGTGTAATGAAATGTTGTCTGGACAGCACTGAAAGAT 2757
QY LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db 2758 AAGAAAGACTAGAGTGTCTGGAGAGTCTGCTACTGAAGCAATAGAAAATCTCCGAAT 2817
QY IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db 2818 GTTGTCTTTCTTGACTCAGGAGCAGAGATTGGAACATATATGATCATCAGAGTTTGCAGGTA 2877
QY GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db 2878 CCATACAGAAACTCTTTGGAAGAAGCACACATCTTGGAAATCAGTTTTCCTTCACGCG 2937
QY AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db 2938 GCAATGATGATTTTCTTCTATGCTGGATGTTTCGGTTTGGAGCCTACTTGTGGACAT 2997
QY GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db 2998 AGTCTCATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTGTCTTGTGGTGCATG 3057
QY AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db 3058 GCCGTGGGCAAGTCAGTTTCATTTGCTCTCTGACTATGCCAAAGCCAAAGPATCAGAGCC 3117
QY HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db 3118 CACATCATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGAGGCCTA 3177
QY LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCys 827
```

```
Db 3178 AAGCCGAACACATTTGGAAGAAATGTCACTTAATGAAGTTGTATTCAACTATCCACC 3237
QY ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db 3238 CGACTGGACATCCAGTGTCTTCAGGGGCTGAGCTTGAAGTGAAGAGGCCACAGCGCTG 3297
QY AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db 3298 GCCTCGTGGCGAGCAGTGGCTGTGGGAAGAGCACCGGTGGTCCAGCTCTCTGGAGCGGTT 3357
QY TyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGlnLeuLeuAsnVal 887
Db 3358 TATGACCCCTTGGCGGGAAAGTGTCTGTACGGCAAGAAATAAAGCACTGAATGTT 3417
QY GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db 3418 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGCAGC 3477
QY IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db 3478 ATTAGTGAGAACATTCCTTATGGAGACAACAGCCGGTGTGTACAGGAAGAGATCGTG 3537
QY GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlnGlyLeuProGluLysTyrAsn 947
Db 3538 AGGCAGCCCAAGAGGCCAATATACAGCCCTTCATCGAGTCACTGCCTAATAAATATAGC 3597
QY ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db 3598 ACCAGATAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAACGCATTGCCATA 3657
QY AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db 3658 GCTCGTGCCTTGTAGACAGCTCATATTTGCTTTTGATGAAGCCACATCAGCTCTG 3717
QY AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db 3718 GATACAGAAGTGAAAGGTTGTCCAAGAGCCCTCGCAAGCCAGAGAGCCCGTACC 3777
QY CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db 3778 TGATTTGCTATGCTCACCGCTGTCCACCATCCAGATTCAGACTTTAATAGTGTGTTT 3837
QY HisAsnGlyLysLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db 3838 CAGATGGCAGATCAAGAGCACCGCACACATCAGCAGCTGTGGCACAGAAAGGCATC 3897
QY TyrPheLysLeuValAsnAlaGlnSer 1056
Db 3898 TATTTTCAATGGTGTCCAGGCT 3924
RESULT 14
AAF86128
ID AAF86128 standard; cDNA; 4195 BP.
XX
AC AAF86128;
XX
DT 25-JUN-2001 (first entry)
XX
DE Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW efflux pump; ss.
OS Macaca fascicularis.
XX
FH Key Location/Qualifiers
FT CDS 100..3951
FT /tag= a
FT /product= "PGP"
FT /note= "P-glycoprotein"
FT misc_feature 376..384
FT /tag= b
FT /note= "Insertion of 9 nucleotides relative to PGP"
```



QY 469 SerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysPro 488  
DB 2167 GCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACT 2226  
QY 489 GluTrpProPheValValLeuGluThrLeuAlaSerValLeuAenGlyThrValHisPro 508  
DB 2227 GAGTGGCCCTATTTCCTTGTGTGTATTTCGTCATTAATAATGGAGTCTGCAACCA 2286  
QY 509 ValPheSerIlePheAlaLysIleThrMetPhe--GlyAenAenAspLysThr 527  
DB 2287 GCATTTGCAGTATATTTTCAAGATATTAGGCAATTTTACAGAAATGATGATGCCGAA 2346  
QY 528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547  
DB 2347 ACAAAACGACAGAAATAGTAACCTTGTTCCTACTATTGTTCTAGCTCTTGCATTTGTTCT 2406  
QY 548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567  
DB 2407 TTTATTACATTTTTCCTTCAGGCTTCACATTTGGCAAAAGCTGGAGATCTCTCAACCAAG 2466  
QY 568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587  
DB 2467 CGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGCTGGTTTGATGAC 2526  
QY 588 LysGluAenSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607  
DB 2527 CCTAAAACACCACTGGAGCATTTGACTACAGGCTCGCCCAATGATGCTGCTCAAGTTAA 2586  
QY 608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAenAlaThrAenMetGlyLeuSer 627  
DB 2587 GGGGCTATAGTTTCCAGGCTGTGTATAATTTACCCAGATATAGCAAAATCTTGGGACAGGA 2646  
QY 628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647  
DB 2647 ATAATTATATCTTAATCATGTTTGGCACTGACACTGTGTACTCTTAGCAATGTGATCCC 2706  
QY 648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAenLysAsp 667  
DB 2707 ATCATTTGCAATACGAGGAGTTGTGAAATGAAATGTTGTCTGGACACGACTGAAAGAT 2766  
QY 668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThr 687  
DB 2767 AAGAAAGAACTAGAGGTGTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAAC 2826  
QY 688 IleValSerLeuThrArgGluLysAlaPheGluGluMetTyrGluGluMetLeuGlnThr 707  
DB 2827 GTTGTTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATGATGATCAGAGTTTGCAGGTA 2886  
QY 708 GlnHisArgAenThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727  
DB 2887 CCATACAGAAACTCTTTGAGAAAGCACACATCTTTGGAATCATCGTTTCTTCACGCGAG 2946  
QY 728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleGlnAla 747  
DB 2947 GCAATGATGATTTTCTTCTATGCTGGATGTTTCCGGTTTGGAGCTACTTGGTGACACAT 3006  
QY 748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767  
DB 3007 AGTCTCATGAGCTTTGAGGATGTTCTGTAGTATTTCAGCTGTGCTTTGTGTGGCATG 3066  
QY 768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787  
DB 3067 GCCGTGGGCAAGTCAGTTCTATTGCTCTGACTATGCCAAAGCAATATCAGCAGCC 3126  
QY 788 HisLeuPheAlaLeuLeuGluLysLysProAenIleAspSerArgSerGlnGluGlyLys 807  
DB 3127 CACATCATCATGATCATTTGAAAAAACCCTTTTGTGATGACAGCTACAGCAGAGGCCCTA 3186  
QY 808 LysProAspThrCysGluGlyAenLeuGluPheArgGluValSerPheTyrProCys 827  
DB 3187 AAGCCGAACACATTTGGAAGGAATGTACACATTTAATGAAGTTGTATTCACTATCCACC 3246

QY 828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847  
DB 3247 CGACTGGACATCCAGTGTCTCAGGGCTGAGCTTGGAAAGTGAAGAGGCCAGACGCTG 3306  
QY 848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867  
DB 3307 GCCCTGTGGCAGCAGTGGCTGTGGAAAGACGCGTGTCCAGCTCTCTGGAGCGGTTTC 3366  
QY 868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAenVal 887  
DB 3367 TATGACCCCTTGGCGGGAAGTGTCTGTGACGCAAAAGAAATAAGCAACTGATGTT 3426  
QY 888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAenCysSer 907  
DB 3427 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCAGC 3486  
QY 908 IleAlaGluAenIleAlaTyrGlyAspAenSerArgValValProLeuAspGluIleLys 927  
DB 3487 ATTAGTGAGAACTTGCCTATGGAGACAAACGCGGTGTGTCCAGGAAGAGATCGTG 3546  
QY 928 GluAlaAlaAenAlaAlaAenIleHisSerPheIleGluGlyLeuProGluLysTyrAen 947  
DB 3547 AGGCGACGCCAAGGAGCCAAATATACACGCTTCATCGACTCACTGCTAATAAATATAGC 3606  
QY 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967  
DB 3607 ACCAGAGTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAAACAACGCATTGCCATA 3666  
QY 968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987  
DB 3667 GCTCGTGCCCTTGTAGACAGCTCATATTTTGTCTTTCGATGAAAGCCACATCAGCTCG 3726  
QY 988 AspAenAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007  
DB 3727 GATACAGAAAGTGAAGAGTTGTCCAGAGGCCCTGGACAAAGCCAGAGGCCCTACC 3786  
QY 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027  
DB 3787 TGCAATGTGATTTGCTCACGCTGTCCACCATCCAGATCCAGACTTAATAGTGTGTTT 3846  
QY 1028 HisAenGlyLysIleLysGluGlnGlyThrHisGlnLeuLeuLeuArgAenArgAspIle 1047  
DB 3847 CAGAATGGCAGAGTCAAGGAGCACGCGCACACATCAGCAGCTGTGGCACAGAAAGCATC 3906  
QY 1048 TyrPheLysLeuValAenAlaGlnSer 1056  
DB 3907 TATTTTCAATGTCAGTGTCCAGGCT 3933  
RESULT 15  
AAV6533  
ID AAV6533 standard; cDNA; 4264 BP.  
XX  
AC AAV6533;  
XX  
DT 08-JAN-1999 (first entry)  
XX  
DE Mutated human P-glycoprotein gene designated mdr-delta-F335.  
XX mdr1-delta-F335; cyclosporin-resistant deletion mutant; human;  
KW P-glycoprotein; cyclosporin; drug efflux;  
KW modulator-resistant multidrug resistance; cancer therapy; ss.  
OS Homo sapiens.  
XX  
PN US5830697-A.  
XX  
PD 03-NOV-1998.  
XX  
PF 21-JAN-1997; 97US-0784649.  
XX  
PR 21-JAN-1997; 97US-0784649.  
XX  
PA (STRD ) UNIV LELAND STANFORD JUNIOR.



XX Chen G, Sikic BI;  
 PI WPI; 1998-609244/51.  
 XX  
 XX DNA encoding mutant P-glycoprotein polypeptide - useful for  
 PT producing recombinant polypeptide  
 XX  
 XX Disclosure; Columns 15-20; 17pp; English.  
 XX  
 CC The present sequence represents mdrl-delta-F335, a gene encoding a  
 CC cyclosporin-resistant deletion mutant of human P-glycoprotein lacking  
 CC Phe335. The recombinant protein can be used to produce antibodies. It  
 CC can also be used in drug screening assays to study the effects of  
 CC modulators such as cyclosporin on P-glycoprotein-mediated drug efflux.  
 CC The nucleic acid can be used to transfer modulator-resistant multidrug  
 CC resistance into transfected cells, e.g. haematopoietic stem cells for  
 CC use during cancer therapy.  
 XX  
 SQ Sequence 4264 BP; 1263 A; 807 C; 1041 G; 1153 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 2,18e-282 Length: 4264  
 Score: 3100.50 Matches: 595  
 Percent Similarity: 75.54% Conservative: 211  
 Best Local Similarity: 55.76% Mismatches: 218  
 Query Match: 58.17% Indels: 43  
 DB: 19 Gaps: 6

US-09-873-409-4 (1-1058) x AAV66533 (1-4264)

QY 4 SerLeuThrSerLysGluLeuSerAlaThrSerLysAlaGlyAlaValaGluVal 23  
 DB TCATTACTGATAAAGAACTCTTAGCTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTGC 909  
 QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
 DB 910 TTGGCAGCAATTAGAACTGTGATTCATTCATTCGAGGACAAAAGAAAGAACTTGAA- 963  
 QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
 DB 963 ----- 963  
 QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysPheGlyIleLys 83  
 DB 964 -----AGTACAAACAAAATTTAGAAAGCTAAAGAAATGGGATTAAG 1008  
 QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
 DB 1009 AAAGCTATTACAGCAATATTCTTATAGTGTCTCTTCCTGCTGTATCTATGATCTTAT 1068  
 QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
 DB 1069 GCTCTGGCCTTCTGTATGGGACCACTTGTCTCTCTCAGGGAA-----TATTCTATT 1122  
 QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
 DB 1123 GGCAAGTACTCACTGATTTCTTCTGTATTAATGGGCTTTTAGTGTGGACAGCA 1182  
 QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
 DB 1183 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGCAGCTTATGAATCTTCAAGATA 1242  
 QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
 DB 1243 ATTGATAATAGCCCAAGTATTGACAGCTATTTCGAAGATGGGCGCAACCAACAGATAATT 1302  
 QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
 DB 1303 AAGGGAATTTGGAAATTCAGAAATGTTCACTTCAGTACCACCTCCGAAAGAGTTAAG 1362  
 QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValaLeuValGlyLeu 223

DB 1363 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGCGACAGCGTGGCCCTGGTTGGAAC 1422  
 QY 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 243  
 DB 1423 AGTGGCTGTGGGAGAGACACACAGTCCAGCTGATCGAGGGCTCTATGACCCACAGAG 1482  
 QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
 DB 1483 GGGATGTCAGTGTTCATGACACAGGATATTAGACCAATAATGAAGGTTTCTACGGAA 1542  
 QY 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
 DB 1543 ATCATTTGGTGTGGTGTAGTCAGGAACCTGTATTGTTGCCACCCAGTAGCTGAAACATT 1602  
 QY 284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
 DB 1603 CGCTATGGCGGTGAAATGTCCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAT 1662  
 QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
 DB 1663 GCCTATGACTTTTATCATGAACCTGCCTCATAAATTTGACACCTCGTGTGGAGAGAGGG 1722  
 QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValaArgAsn 343  
 DB 1723 GCCCAGTGTAGTGGTGGCAGAGCAGAGATGCCCATTTGCACGTCCCTGGTTCCGCAAC 1782  
 QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
 DB 1783 CCCAAGATCTCTCTGTGATGAGGCCAGCTGACGCTTGGACACAGAAAGCAGAGCAGTG 1842  
 QY 364 ValGlnAlaAlaLeuLysAlaSerLysGlyArgThrThrIleValValaIleHisArg 383  
 DB 1843 GTTCAGTGTGCTGTGATGAAGGCCAGAAAAGGTGCGACCACTTGTGTAGTAGCTCATCGT 1902  
 QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
 DB 1903 TTGCTCTACAGTTCGTATGCTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1962  
 QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
 DB 1963 AAAGGAATCATGATGAATCATGAAGAAAGAAAGGCAATTTACTTCAAACTTGTGCACATG 2022  
 QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
 DB 2023 CAGACAGCAGGAATAAGTGAATTTAGAAAATGACGCTGATGAATCCAAAAGTGAAT 2082  
 QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
 DB 2083 GATGCTTGGAAATGTCTTCAAAATGATTCAAGATCCAGTCTAATAAGAAAAGATCAACT 2142  
 QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 470  
 DB 2143 CGTAGAGTGTGCTGGATCAAGCCCAAGCAGAAAGCTTAGTACCAAGAGGCTCTG 2202  
 QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490  
 DB 2203 GATGAAGTATACCTCCAGTTCTTTGGAGGATTAAGCTAAATTTAACTGAATGG 2262  
 QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
 DB 2263 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2322  
 QY 511 SerIlePheAlaValIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
 DB 2323 GCAATATATTTTCAAGATTAAGGGTTTTTACAGAAATGATGATGATCTGAACAAAA 2382  
 QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
 DB 2383 CGACAGATAGTAACCTGTTTCTCACTATTGTTCTAGCCCTTGAATTAATTTCTTTTATT 2442  
 QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
 DB 2443 ACATTTTCTCAGGGTTTTCACATTTTGGCAAGCTGGAGAGATCTCTCAACCAAGCGGCTC 2502



```
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
DB 2503 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2562
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 609
DB 2563 AACACCACTGGAGCATGTGATACAGGCTCGCAATGATGCTCAAGTTAAAGGGGCT 2622
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
DB 2623 ATAGGTTCCAGGCTGCTGTAAATTCATCCAGATATACCAATCTTGGCAGAGNAAT 2682
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
DB 2683 ATATCCTTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATGTACCCATCAT 2742
QY 650 AlaValThrGlyMetIleGlyThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
DB 2743 GCAATAGCAGGAGTTGTTGAAATGAAATGTTCTCGCAAGCACCTGAAAGATAAGAAA 2802
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
DB 2803 GAACAGAGGCTGCTGGAGAGATCGCTACTGAGCAATAGAAAACCTTCCGAAACGGTGT 2862
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
DB 2863 TCCTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
DB 2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTAATCATTTTCTTCCACCCAGGCAATG 2982
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
DB 2983 ATGTATTTTTCCTATGCTGTGATGTTCCGGTTTGGAGCCCTACTGTGGGCACATAAACTC 3042
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
DB 3043 ATGAGCTTTGAGGAGTGTCTGTAGTATTTTACGCTGTGCTTGTGTCCTGATGGCCGTG 3102
QY 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
DB 3103 GGGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAAGCCAAAATATATCAGCAGGCCACATC 3162
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
DB 3163 ATCATGATCATTTGAAAATAACCCCTTTGATTTGACAGCTACAGCAGGAAGCCCTAATGCCG 3222
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 829
DB 3223 AACACATTGGAGGAATGTACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3282
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
DB 3283 GACATCCAGTGTCTTCAGGGACTGAGGCTGGAGGTGAAGAGGCCAGACGCTGGCTCTG 3342
QY 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
DB 3343 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGTGCTCCAGCTCTCTGGAGCGGTTCTACGAC 3402
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
DB 3403 CCCTTGGCAGGGAAGTGTCTGTGATGGCAAGAAATAAGCCACTGANTTTCAGTGG 3462
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
DB 3463 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCATCTCTGTTTACTGTCAGCATGCT 3522
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
DB 3523 GAGAACATTCCTATGGAGACACAGCCGGTGGTGTACAGGAAGAGATTTGTGAGGGCA 3582
```

```
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
DB 3583 GCAAAGGAGGCCCAACATACATGCTTTCATCGAGTCACTGCTTAATAAATATAGCACTAAA 3642
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
DB 3643 GTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAAACGCAATGGCCATAGCTCGT 3702
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
DB 3703 GCCCTTTGTAGACAGCCTCATATTTTGTGTTGGATGAAGCCAGCTCAGCTCTGGATACA 3762
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
DB 3763 GAAAGTGAAAGGTTTGTCCAAGAAAGCCCTGGACAAAGCCAGAGAGGCCACCTGCATT 3822
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
DB 3823 GTGATTTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGAAT 3882
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049
DB 3883 GGCAGAGTCAGGAGCATGGCAGCATCAGCAGCTGCTGGCAGACAGAAAGGCATCTATTTT 3942
QY 1050 LysLeuValAsnAlaGlnSer 1056
DB 3943 TCAATGGTCAGTGTCCAGGCT 3963
```

Search completed: March 30, 2003, 03:12:46  
Job time : 1302.02 secs

**THIS PAGE BLANK (USPTO)**



; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-784-649A-1

## Alignment Scores:

Pred. No.: 0 Length: 4264  
Score: 3100.50 Matches: 595  
Percent Similarity: 75.54% Conservative: 211  
Best Local Similarity: 55.76% Mismatches: 218  
Query Match: 58.17% Indels: 43  
DB: 2 Gaps: 6

US-09-873-409-4 (1-1058) x US-08-784-649A-1 (1-4264)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23  
Db 850 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 909  
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
Db 910 TTGCAGCAATTAGAAGTGTGATTCATTTGGAGGACAAAGAAAGAACTTGAA----- 963  
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
Db 963 ----- 963  
QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 964 -----AGGTACACCAAAATTTAGAAGACTAAAGATTGGGATAAG 1008  
QY 84 ArgThrIleAlaSerLysValSerGluValAlaValTyrPhePheMetAsnGlyThrTyr 103  
Db 1009 AAAGCTATTACAGCAATATTTCTAGGTGCTCTTCTCGTGTATGATCATCTTAT 1068  
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
Db 1069 GCTCTGGCCCTTCGTATGGGACCACTTGTCTCTCTCAGGGAA-----TATCTATT 1122  
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
Db 1123 GGCAAGTACTCACTGTATTCTTTCTGTATTAATTTGGGCTTTTAGTGTGGACGCA 1182  
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 1183 TCTCCAGCATTTGAAGCATTTGCAATTCGAAGAGCAGCTTATGAAATCTTCAAGATA 1242  
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 1243 ATTGATAAAGCCCAAGTATTGACAGCTATTGGAAGAGTGGGCACAAACCCAGATAATT 1302  
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
Db 1303 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 1362  
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1363 ATCTTGAAGGCCCTGAAACCTGAAGCTGAGAGTGGGAGCAGCGTGGCCCTGGTTGGAAC 1422  
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
Db 1423 AGTGGCTGTGGAGAGACACACACAGTCCAGCTGATGACAGGCTCTATGACCCACAGAG 1482  
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1483 GGGATGGTCAGTGTGATGGACAGGATATTAGGACCATTAATGTAAGGTTTCTACGGGA 1542  
QY 264 HistLeuGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1543 ATCAATGGTGGGAGTTCAGAACCTGTATGTTGCCACCAAGTAGCTGAAACATT 1602  
QY 284 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 303

Db 1603 CGCTATGGCCGTGAAATGTCCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662  
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
Db 1663 GCCTATGACTTTATCATGAACCTGCTCATAAATTTGACACCTGTTGGAGAGAGGG 1722  
QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1723 GCCCAGTTGAGTGGGCGAGAGCAGAGATCGCCATTGCACGTGCCCTGTTGCGAAC 1782  
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1783 CCCAAGATCCTCTGCTGATGAGGCCAGCTCAGCCCTTGGACACAGAAAGCGAAGCAGTG 1842  
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 1843 GTTCAGGTGCTCTGATAGGCCAGAAAGGTCGACACCACTTGTATGATGCTCATCGT 1902  
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1903 TTGCTACAGTTGCTATGCTGACGTCTCGCTGTTTCGATGATGGAGTCATTGTGGAG 1962  
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 1963 AAAGGAAATCATGATGAATCATGAAAGAGAAAGGCACTTACTTCAAACTTGTCACAATG 2022  
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
Db 2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGACGCTGATGAATCCAAAGTGAAATT 2082  
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 2083 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCCAGTCTAATAAGAAAAAGCAACT 2142  
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470  
Db 2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGCAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2202  
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuLysProGluTrp 490  
Db 2203 GATGAAGTATACCTCCAGTTCTCTTTGGAGGATATGAGCTAAATTTAACTGAATGG 2262  
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2263 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2322  
QY 511 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2323 GCAATAATATTTCAAAGATATATAGGGTTTTTACAAGAATTGATGATCTCTCAAAACAAA 2382  
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2383 CGACAGATAGTAACTGTTTTCCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2442  
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2443 ACATTTTCTTCAGGGTTTCATTTTGGCAAGCTGGAGAGATCTCACCACAGCGGCTC 2502  
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589  
Db 2503 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2562  
QY 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2563 AACACCACTGGAGCATGATACACAGCTCGCAATGATGCTGCTCAAGTTAAAGGGGCT 2622  
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2623 ATAGTTTCCAGGCTTCTGTAATATCCAGAAATATAGCAATCTTGGACAGGAATAATT 2682  
QY 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649

Db 2683 ATATCTCTCATCTATGGTTGGCACTAAACACTGTTACTCTTAGCAATTGTACCCATCATTT 2742  
Qy 650 AlaValThrGlyMetLeuGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2743 GCAATAGCAGGAGTCTGTGAATGAAATATGTTCTGTGGCAACGACTGAAAGATAAGAA 2802  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnLysArgThrIleVal 689  
Db 2803 GAACTAGAAAGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2862  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLysGlnThrGlnHis 709  
Db 2863 TCCTTACTCAGAGGAGGAGGAGTTTGACATATATGTATGCTCAGAGTTTGAGGTACCATAC 2922  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 2923 AGAACTCTTTGAGGAAAGCACACATCTTTTGAATATACATTTTCTTCCACCCAGGCAATG 2982  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuLeuGlnAlaGlyArg 749  
Db 2983 ATGTATTTTTTCCATCTGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAATC 3042  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 3043 ATGAGCTTTCAGGATGTTCTGTAGTATTTTACGTGTTCTTGTGGTGCATGGCGGTG 3102  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 3103 GGGCAAGTCAGTTTCATTTGCTCTGCTGACTATGCAAGCCAAATATATCAGCAGCCACATC 3162  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3163 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGCAGGAGGCGCTTAATGCCG 3222  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 3223 AACACATTGGAAGGAATGTACATTTGTTGTAAGTTGATTAACCTATCCACCCGACCG 3282  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3283 GACATCCAGTGTCTCAGGAGCTGAGCTGGAGTGAAGAGGCCAGAGCGCTGGCTCTG 3342  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3343 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTCGAGCGGTCTACGAC 3402  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889  
Db 3403 CCTTGGCAGGGAAGTGTCTGTGATGGCAAGAAATAAAGCGACTGAATGTTCACTGG 3462  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3463 CTCGAGCACACCTGGCATCTGTGCCAGAGCCCATCTGTTGACTGCAGCATGCT 3522  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3523 GAGAAATTCCTATGTGAGACAAACAGCGCGGTGGTGTACAGGAAGAGATGTGTGAGGCA 3582  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3583 GCAAGAGGAGCCCAACATACATGCTTTCATCGAGTCACTGCTTAATAATATAGCACTAAA 3642  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3643 GTAGGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAGCATGTGCTAGCTCGT 3702  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 3703 GCCCTTGTAGACAGCCTCATATTTTCTTTTGGATGAAGCCAGCTCAGCTCTGGATACA 3762  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 3763 GAAAGTGAAAGGTTGTCCAAGAGGCGCTCGACAAAGCCAGAGAGGCGCCGACCTGCATT 3822

Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleuLeuValValLeuHisAsn 1029  
Db 3823 GTGATTGCTCACGCGCTGCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAGAAT 3882  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 3883 GGCAAGTCAAGAGCATGCGCAGCATCAGACGCTCTGGCACAGAAGCGCATCTATTTT 3942  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 3943 TCAATGTCAGTGTCCAGGCT 3963  
RESULT 2  
US-08-784-649A-5  
; Sequence 5, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,649A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: Reg. No. 5830697 36,677  
; REFERENCE/DOCKET NUMBER: 06037/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5070  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4264 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
US-08-784-649A-5  
Alignment Scores:  
Pred. No.: 0 Length: 4264  
Score: 3100.50 Matches: 595  
Percent Similarity: 75.54% Conservative: 211  
Best Local Similarity: 55.76% Mismatches: 218  
Query Match: 58.17% Indels: 43  
DB: 2 Gaps: 6

US-09-873-409-4 (1-1058) x US-08-784-649A-5 (1-4264)  
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23  
Db 850 TCATTTTACTGATAAAGAACTCTTTAGCGTATGCAAAAAGCTGGAGCAGTAGCTGAAGAGTC 909  
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
Db 910 TTGGCAGCAATTAGAACTGTGATTGTCATTTGGAGGACAAAGAAAGAACTTGAA----- 963

QY 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
Db 963 ----- 963  
QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspPheGlyIleLys 83  
Db 964 -----AGGTACACAAATAATTTAGAAAGCTAAAGAAATGGGATAAG 1008  
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyr 103  
Db 1009 AAGCTATTACAGCCAAATATTTCTAGGTGCTGCTTCTCTGCTATGATCTAT 1068  
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 123  
Db 1069 GCTCTGSCCTTCTGTTATGGACACCTTGCTCTCTCAGGGAA-----TATTCTATT 1122  
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
Db 1123 GGACAAGTACTCACTGTATTCTTTCTGTATTATTTGGGCTTTTAGTGTGGACAGGCA 1182  
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 1183 TCTCCAGCAATTGAAGCATTTGCAATTCGAAGAGGACGCTTATGAATCTTCAAGATA 1242  
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 1243 ATTGATAATAAGCCAACTATTGACAGCTATTTCGAAGAGTGGGCACAAACAGATAATT 1302  
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
Db 1303 AAGGGAAATTTGGAAATTCAGAAATGTTTCACTTCAGTTTACCCTATCTCGAAAGAGTTAAG 1362  
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1363 ATCTTGAGGCGCTGAACCTGAAGTGCAGAGTGGCAGCGTGGCCCTGGTTGGAAC 1422  
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 243  
Db 1423 AGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1482  
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1483 GGGATGCTCAGTGTGTGAGACAGGATATTAGGACCAATAATGTAAAGTTTCTACGGGAA 1542  
QY 264 HistGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1543 ATCAATGGTGTGGGTGAGTCAGAACCTGTATTGTTGCCACCATGATGCTGAAACATT 1602  
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1603 CGCTATGCCCTGAAATGTCCACATGGATGAGATTGAGAAGCTGTCAGGAAGCCAA 1662  
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
Db 1663 GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 1722  
QY 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1723 GCCCAGTGTAGTGTGGGCAAGCAGAGATGCCATTCGACGTGCCCTGGTGGTTCGCAAC 1782  
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1783 CCCAGATCTCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGGAGCAGTG 1842  
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArg 383  
Db 1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAAGGTCCGACCCACCATTTGTGATAGCTCATCGT 1902  
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1903 TTGTCACAGTTGTAAATGCTGACGTCATCGCTGGTTTCCGATGATGAGTCATTGTCGGAG 1962  
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423

Db 1963 AAAGGAAATCATGATCACTCATGAAGAGAAAGGCAATTTACTTCAAACTGTCACAATG 2022  
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
Db 2023 CACACAGCAGGAAATGAAGTTGAATTAGAAAATGCACTGCTAGTAATCCAAAGAGTAAT 2082  
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 2083 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACT 2142  
QY 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470  
Db 2143 CGTAGGAGTGTCCGTGGATCACAGGCCAACAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2202  
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuLysProGluTyr 490  
Db 2203 GATGAAGTATACCTCCAGTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2262  
QY 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2263 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2322  
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2323 GCAATAATATTTTCAAGAGATTAAGGGGTTTTTACAAGAATTGATGATCTGTAACAAAA 2382  
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2383 CGACAGATAGTAGTAACTTGTGTTTCTACTATTGTTTCTAGCCCTTGGAAATTTATCTTTAT 2442  
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2443 ACATTTTCTCTCAGGTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGGCTC 2502  
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589  
Db 2503 CGATACATGGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGTTGATGACCCCTAAA 2562  
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2563 AACACCATGGAGCATGTGATCAGCGCTGCCAATGTAGTGTCTCAAGTTAAAGGGCT 2622  
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2623 ATAGTTTCCAGGCTGCTGTAATACCCAGATATAGCAATCTTGGACAGGAAATAT 2682  
QY 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2683 ATATCTTCACTATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCATCAT 2742  
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2743 GCAATACAGGAGCTGTTGAAATGAAATGTTGCTCTGGACAAAGCTGAAAGATAGAAA 2802  
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2803 GAATAGAGGTCTGGGAGATCGCTACTGAAGCAATAGAAAACTTCCGAAACGTTGTT 2862  
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 2863 TCTTTGACTCAGGACGAGAGATTGAAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922  
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 2923 AGAACTCTTTGAGGAAAGCACACATCTTTTGGAAATTTACATTTCTCTTCCACCCAGGCAATG 2982  
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValatTyrLeuIleGlnAlaGlyArg 749  
Db 2983 ATGTATTTTCTTATGCTGATGTTTCCGTTTGGAGCTTACTTTGGTGGCACAATAAATC 3042  
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769







```
Db 3569 GACATCCAGTGCTTCAGGAGCTAGCCTGGAGGTGAAGAGGCCAGAGCGCTGGCTCTG 3628
Qy 850 ValGlySerSerGlyCysGlyLysThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db 3629 GTGGGAGCAGTGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGTTCTACGAC 3688
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889
Db 3689 CCCTTGGCAGGGAAGTGTCTGTATGGCAAGAAATAAAGGACGTGAATGTTCAGTGG 3748
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3749 CTCGAGCACACCTGGCATCGTGTCCAGAGCCCATCTCTGTTCAGTCGACATGCT 3808
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db 3809 GAGAACATTGCTATGGAGACAAACAGCCGGTGGTGTACAGGAGAGATCGTGAGGGCA 3868
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db 3869 GCAAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCTAATAAATATAGCACTAA 3928
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3929 GTAGGAGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAGCATTCGCTAGTCTGT 3988
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3989 GCCTTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACA 4048
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4049 GAAAGTGAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 4108
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 4109 GTAGTGTCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAGAA 4168
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4169 GGCAGATCAGAGAGATGGCAGCATCAGCAGCTGCTGCGCACAGAAAGCATCTATTT 4228
Qy 1050 LysLeuValAlaGlnInser 1056
Db 4229 TCAATGTCAGTGTCCAGGCT 4249

RESULT 4
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 6 Gaps: 6
```

```
US-09-873-409-4 (1-1058) x 5206352-3 (1-4669)
Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValaGluVal 23
Db 1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195
Qy 24 LeuSerSerIleArgThrValIleAlaPheAArgAlaGlnGluLysGluLeuGlnAArgSer 43
Db 1196 TTGGCAGCAATTAGAACTGTGATTTGGAGGACAAAAGAAAGAACTTGAA----- 1249
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
Qy 64 ValLeu**PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACAACAAAATTTAGAAGAGCTAAAGAATTTGGATAG 1294
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAAGCTATTACAGCCAATATTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCATCTAT 1354
Qy 104 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 1355 GCTCTGGCTTCTGATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1408
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGAAGTACTACTGATATTCTTCTGATTAATTTGGGGCTTTTAGTGTGACAGGCA 1468
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1469 TCTCCAAGCATTTGAACATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1529 ATTGATAAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCACAAACCCAGATAATAT 1588
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1589 AGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAGAGTTAAG 1648
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1649 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGAGCGTGGCCCTGGTTGAAAC 1708
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1709 AGTGGCTGTGGGAAGAGACAAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCAAGAG 1768
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1769 GGATGCTCAGTGTGATGACAGAGATATTAGACCATTAATGTAAGTTTCTACGGGAA 1828
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1829 ATCATTGGTGTGAGTCAAGAACCTGTATTGTTGGCCACCAGCATAGCTGAAACATT 1888
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1889 CGCTATGGCGGTGAAATGTCAACATGATGAGATTTGAGAAAGCTGTCAAGGAGGCAAT 1948
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1949 GCCTATGACTTTATATGAACCTGCTCAATAATTTGACACCTGTTGGAGAGAGGG 2008
Qy 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 2009 GCCCAGTTGAGTGGTGGCAGAGAGATGCCATTTGCATGCTGCCCTGGTTCCGCAAC 2068
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
```

Db 2069 CCCAAGATCTCTCTGGATGAGGCCACCTCAGCCTTGGACACAGAAAACCGAAGCAGTG 2128  
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyValArgThrThrIleValValAlaHisArg 383  
Db 2129 GTTCAGGTGCTCTGGATAGGCCAGAAAAGTTCGGACCACTTGTGATGCTCATCGT 2188  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 2189 TTGTCTACAGTTCGTAATGCTGATGCTCATCGCTGCTTCGATGATGAGTCATTTGGAG 2248  
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 2249 AAAAGGAATCATGATGAACCTCATGAAGAGAAAGGCAATTTACTTCAAACCTGTGCACAA 2308  
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
Db 2309 CAGACACGAGAAATGAAGTTGAATTAAGAAATGCACCTGATGATCCAAAGTGAAT 2368  
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 2369 GATGCTTGGAAATGCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAGATCAACT 2428  
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluLysThrGlnSerLys 470  
Db 2429 CGTAGGAGTGTCCGTGATCACAAAGCCAAAGAGCTTAGTACCAAGAGGCTCTG 2488  
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490  
Db 2489 GATGAAGTATACCTCCAGTTCCTTTGGAGGATTAAGCTTAAATTTAACTGAATGG 2548  
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2549 CCATATTTTGTGTGTGTTTGTGTCATTTTGTGCCATTAATAATGGAGCCCTGCAACGACATTT 2608  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2609 GCAATAATATTTTCAAAGATATATAGGGGTTTTTACAAGAAATTTGATGATCTCTGAACAAAA 2668  
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2669 CGACAGAAATAGTAACTTGTCTTCTACTATTTGTTCTAGCCCTTGAATATTTCTTTTATT 2728  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2729 ACATTTTCTTCAGGTTTCACATTTTGGCAAGCTGGAGAGATCTCCACCAAGCGCTC 2788  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589  
Db 2789 CGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTTGTATGACCCCTAAA 2848  
Qy 590 AsnSerThrGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2849 AACACCACTGGAGCATGATACCAAGCTCCCAATGATGCTCAAGTTAAAGGGCT 2908  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2909 ATAGTTCAGGCTTGTGTAAATACCCAGATATAGCAATCTTGGACAGGAATATT 2968  
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2969 ATATCTTTCATCTATGTTGGCAACTAACACTGTACTCTTAGCAATTTGATCCCATCATT 3028  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 3029 GCAATAGCAGGAGTGTGTGAATGAAATGTTGTGGCAAGCACTGAAGAGATAAGAAA 3088  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 3089 GAACTAGAGGTGTGGAGAGATCGCTACTGAAGCAATAGAAACTTCCGAAACCGTTGTT 3148  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 3149 TCTTTGACTCAGGACGAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATC 3208

Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 3209 AGAAACTCTTTGAGGAAAGCACACATCTTTTGGAAATACATTTTCTTCCACCGGCAATG 3268  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749  
Db 3269 ATGTATTTTCTTATCTGATGATGTTTCCGGTTTGGAGCCTTACTTGTGGCACATAAACTC 3328  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 3329 ATGAGCTTTGAGATGTTCTGTAGTATTTTTCAGCTGTGTCTTTGGTGCATGGCCGTG 3388  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789  
Db 3389 GGGCAAGTCAGTTCATTTGCTCTGACTATGCCAAAGCCAAATATCAGCAGGCCACATC 3448  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3449 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACACGACGGAAGCCCTAATGCCG 3508  
Qy 810 AspThrCysGluGluAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 3509 AACACATTTGAAGGAATGTACATTTGTTGTAAGTTGTATTCAACTATCCACCCGACCG 3568  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3569 GACATCCAGTGTCTCAGGACTGAGCTGGAGTGAAGAGGCCAGACGCTGGCTCTG 3628  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3629 GTGGGACGAGTGGCTGTGGAAAGACACAGTGTGCTCAGCTCTGGAGCGGTCTTACGAC 3688  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889  
Db 3689 CCCTTGGCAGGAAGTGTCTGTGATGGCAAGAAATAAAGCAGCTGAATGTTCACTGG 3748  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3749 CTCGAGCACACCTGGGCATCGTGTCCAGAGGCCCATCTCTGTTGACTGCAGCATTTGCT 3808  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3809 GAGAACATTTCCCTTATGGAGACAAACAGCCGGTGTGTGTACAGGAAGAGATCGTGAGGGCA 3868  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3869 GCAAGAGGCCCAACATACATGCTTCTCATCGAGTCACTGCTTAATAATAATAGCAGCTAAA 3928  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3929 GTAGGACAAAGAACTCAGCTCTCTGTTGGCCAGAAACACGCACTTGCCATAGCTCGT 3988  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 3989 GCCCTTGTAGACAGCTCATATTTTGTCTTTGGATGAAGCCAGAACGCCAGCCCTGCAAT 4048  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 4049 GAAAGTGAAGAGTGTTCGAAGAGCCCTGGACAAAGCCAGAGAACGCCAGCCCTGCAAT 4108  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 4109 GTGATTCCTCAGCCCTGTCCACCATCCAGATGCAGACTTAATAGTGTGTTTTCAGAT 4168  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 4169 GGCAGAGTCAAGAGAGTGGCAGCATCAGCAGCTGTGGCAGAGAAAGGCATCTATTTT 4228  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 4229 TCAATGCTCAGTGTCCAGGCT 4249

## RESULT 5

US-08-793-610-5  
; Sequence 5, Application US/08793610  
; Patent No. 5858744  
; GENERAL INFORMATION:  
; APPLICANT: BAUM, Christopher  
; APPLICANT: STOCKING-HARBERS, Carol  
; APPLICANT: OSTERTAG, Wolfram  
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
; TITLE OF INVENTION: FOR GENE TRANSFER  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikolaïdo, Marmelstein, Murray & Oram LLP  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,610  
; FILING DATE: 07-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 31 973.8  
; FILING DATE: 08-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 03 952.1  
; FILING DATE: 07-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/03175  
; FILING DATE: 10-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berman, Richard J.  
; REGISTRATION NUMBER: 39,105  
; REFERENCE/DOCKET NUMBER: P1614-7007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6505 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA

US-08-793-610-5

Alignment Scores:  
Pred. No.: 0 Length: 6505  
Score: 3100.50 Matches: 595  
Percent Similarity: 75.54% Conservative: 211  
Best Local Similarity: 55.76% Mismatches: 218  
Query Match: 58.17% Indels: 43  
DB: 2 Gaps: 6

US-09-873-409-4 (1-1058) x US-08-793-610-5 (1-6505)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23  
Db 2528 TCATTTACTGATAAGAACTCTTAGCGTATGCAAAAGCTGGACAGTACTGAGAGGTC 2587  
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
Db 2588 TTGGCAGCAATTAGAACTGTGATTGCTATTGGAGGACAAAGAAAGAACTTGAA----- 2641  
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
Db 2641 -----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437

Qy 64 ValLeu\*\*\*PheValArgTyrThrClnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 2642 -----AGTCAACAACAAATTTAGAAAGCTAAAGAAATTGGGATAAG 2686  
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
Db 2687 AAAGCTATTACAGCCAATATTCTATAGTGTCTTCTCTCTCTCTCTCTCTCTCTAT 2746  
Qy 104 GlyLeuAlaPheTyrTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 123  
Db 2747 GCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCTCTCTCTCTCTCTCTCTCTAT 2800  
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
Db 2801 GGCAAGTACTCACTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2860  
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 2861 TCTCCAGCAATTGAAGCAATTTGCAATTCAGAGAGAGAGAGCTTATGAATCTTCAAGATA 2920  
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 2921 ATTGATAATAAGCCAAGTATTGACAGCTATTTCGAAGAGTGGGACACAAACACGATAAT 2980  
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
Db 2981 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTATCCCATCTCGAAAAGAAAGTTAA 3040  
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 3041 ATCTTGAAGGGCTGAACTGAAGTGCAGAGTGGGACAGCGTGGCCCTGTTGGAAAC 3100  
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
Db 3101 AGTGGCTGTGGGAGAGACACACAGTCCAGCTGATGCAGAGGCTCTATCACCCACAGAG 3160  
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 3161 GGGATGGTCAGTGTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3220  
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 283  
Db 3221 ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTTGGCCACCACGATAGCTGAAACATT 3280  
Qy 284 LysTyrGlyValArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 3281 CGCTATGGCGTGAAATGTCAACATGGATGATGAGATGAGAAAGCTCAAGGAAGCCAAT 3340  
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
Db 3341 GCCTATGACTTTATCATGAACTGCTCTATAAATTTGACACCTCTGGTTGGAGAGAGGG 3400  
Qy 324 AlaGlnMetSerGlyGlyClnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 3401 GCCCAGTTGAGTGGTGGGAGAGAGAGAGATCGCAATTCAGCTGCTCTGGTTCCCAAC 3460  
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 3461 CCCAGATCTCTCTGCTGATGAGCCACGTCAGCTCTGCACACAGAAAGCGAGCAGTG 3520  
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 3521 GTTCAGTGGCTCTGATTAAGGCCAGAAAGGTTCGAGCACCACTTGTGTAGTCTATCGT 3580  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 3581 TTGCTACAGTTGCTGAATGCTGACGTCATCGCTGGTTCGATGATGGAGTCATTGTGGAG 3640  
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 3641 AAAGGAATCATGATGAACTCATGAAGAAAGGCAATTTACTTCAAACTTGTCAACATG 3700  
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437

Db 3701 CACACAGCAGGAAATGAAGTTGAATAGAAATGCAGCTGATGAATCCAAAAGTGAAT 3760  
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal 452  
Db 3761 GATGCGCTTGGAATGCTTCAATAGATTCAGAGTCCAGTCTAATAAGAAAGAGTCAACT 3820  
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 470  
Db 3821 CTGAGGAGTGTGCGTGGATCACAAGCCACAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3880  
Qy 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTTP 490  
Db 3881 GATGAAAGTATACCTCCAGTTCTCTTTGAGGATTATGAAGCTAAATTTAACTGAATGG 3940  
Qy 491 ProPheValValLeuGluThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 3941 CCTATTCTTGTGTGTGATTGTTGGCCATTATAAATGGAGCCTCAACCAAGCATTT 4000  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 4001 GCAATATATTTCAAGATATATAGGGTTTTACAGAGATTGATGATCCTGAAACAAA 4060  
Qy 530 LysHisAspAlaGluIleThrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 4061 CGACAGAAATAGTAACCTGTTCTCACTATTGTTCTAGCCCTTGGAAATTTCTTTTATT 4120  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 4121 ACATTTTCTTCAGGGTTCACATTTGGCAAGCTGGAGAGTCTTCACCAAGCGGCTC 4180  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTTPPheAspGluLysGlu 589  
Db 4181 CGATACATGTTTCCGATCCATGCTCAGACAGAGTGTGAGTTGGTTGATGCCCTAAA 4240  
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 4241 AACACCACTGGAGCATTTGACTTACCAGGCTCGCCAATGATGCTCAAGTTAAAGGGCT 4300  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 4301 ATAGGTTCCAGGCTTGCTGTAATTTACCAGAAATATAGCAATCTTGGGACAGGAATATT 4360  
Qy 630 IleSerPheIleTyrGlyTTPGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 4361 ATATCTTTCATATGTTGGCACTAACACTGTTACTCTTAGCAATGTACCATCATTT 4420  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 4421 GCAATAGCAGGAGTGTGAAATGAAATGTTGTCTGGACAGCACTGAAAGATAAGAAA 4480  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 4481 GAACCTAAGGTGTGGAAGATCGCTACTGAACCAATAGAAAACCTCCGAAACCGTTGTT 4540  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 4541 TCTTTGACTAGGAGCAGAGTGTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4600  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 4601 AGAAACTCTTTGAGAAAGCACATCTTTGGAATTACATTTCTCTTCAACCCAGCAATG 4660  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749  
Db 4661 ATGATTTTCTTATGTTGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACAATAACTC 4720  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 4721 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGCTTTGGTGGCATGCGCGTG 4780  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789

Db 4781 GGGCAAGTCAGTTCAATTGCTCTGACTATGCAAGCCAAAGCAAAATATCAGCAGCCACATC 4840  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 4841 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGCAGCGAAGCCCTAATGCCG 4900  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 4901 AACACATTGGAAGAAATGTCATTTGTTGAAGTTGTTATTCAACTATCCACCCGACCG 4960  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 4961 GATCCTCCAGCTGCTCAGGAGCTGAGCTGGAGTGAAAGAGGCGCAGACCTGGCTCTG 5020  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 5021 GTGGGACAGAGTGGCTGTGGAGAGACACAGTGTCTCCAGCTCTCTGGAGCGGTTCTACGAC 5080  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTTP 889  
Db 5081 CCTTTGGCAGGAAAGTGTCTGTTGATGGCAAGAAATAAAGCGACTGAATGTTCACTGG 5140  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 5141 CTCGAGACACCTGGGCATCGTGTCCAGAGCCCATCTGTTTGCATCGACATGCT 5200  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929  
Db 5201 GAGAACATTTGCTATGGACAGAACACAGCGGTGGTGTACAGGAAGAGATCGTGAGGGCA 5260  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 5261 GCAAGAGGAGCCAAACATACATGCTTTCATCGAGTCACTGCCTAATAATATAGCACTAAA 5320  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 5321 GTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAAAGCAATGCCATGCTCGT 5380  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 5381 GCGCTTTGATAGAGCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACA 5440  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 5441 GAAAGTGAAGAGTTTCTCCAAGAGCCCTCGGACAAAGCCAGAGAGCGCCACCTGCATT 5500  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 5501 GTGATTGCTCAGCGCTGCTCCACCATCCAGATGCAGACTTAATAGTGTGTTTCAGAAAT 5560  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPhe 1049  
Db 5561 GGCAGAGTCAAGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCACTATTTT 5620  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 5621 TCAATGTCAGTGTCCAGGCT 5641

## RESULT 6

US-08-793-610-6  
; Sequence 6, Application US/08793610  
; Patent No. 5858744  
; GENERAL INFORMATION:  
; APPLICANT: BAUM, Christopher  
; APPLICANT: STOCKING-HARRIS, Carol  
; APPLICANT: OSTERTAG, Wolfram  
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington

STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005-5701  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/793,610  
 FILING DATE: 07-MAR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE P 44 31 973.8  
 FILING DATE: 08-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 195 03 952.1  
 FILING DATE: 07-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP95/03175  
 FILING DATE: 10-AUG-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Berman, Richard J.  
 REGISTRATION NUMBER: 39,105  
 REFERENCE/DOCKET NUMBER: P1614-7007  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)638-5000  
 TELEFAX: (202)638-4810  
 INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 9318 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA

US-08-793-610-6

#### Alignment Scores:

Pred. No.:	0	Length:	9318
Score:	3100.50	Matches:	595
Percent Similarity:	75.54%	Conservative:	211
Best Local Similarity:	55.76%	Mismatches:	218
Query Match:	58.17%	Indels:	43
DB:	2	Gaps:	6

US-09-873-409-4 (1-1058) x US-08-793-610-6 (1-9318)

QY	4	SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal	23
DB	2487	TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC	2546
QY	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
DB	2547	TTGGCAGCAATTAGAAGCTGTGATTTGGAGGACAAAGAAAGAACTTGAA-----	2600
QY	44	PheLeuLeuAsnIleThrArgTyrAlaTyrPheProGlnTyrLeuLeuSerCys	63
DB	2600	-----	2600
QY	64	ValLeu***PheValArgTyrThrGlnAsnLeuLysAspPheGlyIleLys	83
DB	2601	-----AGGTCAACAAAAATTTAGAAGAACTTAAAGAAATTTGGATAAAG	2645
QY	84	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThrTyr	103
DB	2646	AAAGCTATTACAGCCCAATATTTATAGGTGCTCTTCTCGTGTATGATCATCTTAT	2705
QY	104	GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	123
DB	2706	GCTCTGGCCCTTCTGTTATGGACACCTTGGTCTCTCAGGGNA-----TATCTATT	2759
QY	124	GlyThrValLeuAlaValPheSerValIleHisSerTyrCysIleGlyAlaAla	143
DB	2760	GGACAAGTACTCACTGTATTCTTCTGTATTAAATTGGGGCTTTTAGTGTGGACAGCA	2819

QY	144	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	163
DB	2820	TCTCCAAGCATTGAAGCATTTCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA	2879
QY	164	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	183
DB	2880	ATTGATAATAAGCCAAGTATTACAGCTATTGAAAGGTGGGCACAAACAGATAATATT	2939
QY	184	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
DB	2940	AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCCATCTCGAAAGAGTTAG	2999
QY	204	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
DB	3000	ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGCAGACGGTGGCTGGTTGGAAC	3059
QY	224	AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	243
DB	3060	AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGACAGGCTCTATGCCCCACAGAG	3119
QY	244	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	263
DB	3120	GGATGCTCAGTGTGATGACAGGATATTAGCACCATTAATGTAAAGTTTCTACGGGA	3179
QY	264	HisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle	283
DB	3180	ATCATTTGGTGTGGTGAAGTCAAGAACCTGTATTGTTGCCACCACGATAGCTGAAACAT	3239
QY	284	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
DB	3240	CGCTATGGCGGTGAATAATGTCACCATGATGATGAGAAAGCTGTCAAGGAAGCAAT	3299
QY	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323
DB	3300	GCCTATGACTTTATCATGAAGACTGCTCTATAATTTGACACCTGTTGGAGAGAGGG	3359
QY	324	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
DB	3360	GCCAGTTGAGTGGTGGCAGAGCAGAGGATCGCCATTCGACGTGCCCTGGTTCCGCAAC	3419
QY	344	ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
DB	3420	CCCAAGATCTCTCTGCTGATGAGGCCAGCTCAGCTTGGACACAGAAAGCGAGCAGT	3479
QY	364	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArg	383
DB	3480	GTTCAAGTGGCTCTGATAGGCCAGAAAGGTGGACACCATTTGTATGATCTCATCGT	3539
QY	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
DB	3540	TTGCTTACAGTTCTGTAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGTGGAG	3599
QY	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	423
DB	3600	AAAGGAAATCATGATCAACTCATGAAAGAGAAAGGCACTTTACTTCAAACTGTCACAATG	3659
QY	424	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	437
DB	3660	CACACAGCAGGAAATGAAGTTGAATTAGAAATGACGAGTGAATCCAAAGAGTGAATT	3719
QY	438	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	452
DB	3720	GATGCTTGGAAATGCTTCAAAATGATTACAGATCAGCTCATTAAGAAAAAGATCAACT	3779
QY	453	---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys	470
DB	3780	CGTAGGAGTGTCCGTGGATCACAAGCCCAAGCAGAAAAGCTTAGTACCAAGAGGCTCTG	3839
QY	471	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr	490
DB	3840	GATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTTAGAACTTAAATTAACCTGAATGG	3899







## ATTORNEY/AGENT INFORMATION:

NAME: King, William T  
REGISTRATION NUMBER: 30,954  
REFERENCE/DOCKET NUMBER: GP50008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5015  
TELEFAX: 610-270-5090  
TELEX:

## INFORMATION FOR SEQ ID NO: 1:

## SEQUENCE CHARACTERISTICS:

LENGTH: 4233 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA

US-09-120-513-1

## Alignment Scores:

Pred. No.: 0 Length: 4233  
Score: 3086.00 Matches: 601  
Percent Similarity: 74.74% Conservative: 198  
Best Local Similarity: 56.22% Mismatches: 224  
Query Match: 57.90% Indels: 46  
DB: 3 Gaps: 7

US-09-873-409-4 (1-1058) x US-09-120-513-1 (1-4233)

Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20  
Db 722 GTACTGACTTCTTACTATAAGAACTCCAGCTTATGCGAAGCTGAGGAGTTGCC 781  
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40  
Db 782 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAAAAGAGGAATT 841  
Qy 41 GlnArgSerPheLeuLeuAlaIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60  
Db 842 GAA----- 844  
Qy 61 LeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80  
Db 845 -----AGTACATATAAATTTAGAAAGCTAAAGAGTT 880  
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsn 100  
Db 881 GGCATTAAGAAACCCATCAGCGCCAACTTCCATAGTATTGCTACCTGTTGGTCTAT 940  
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120  
Db 941 GCGTCTTATGCACTGCGCATTTCTGTATGGGACCTCTTGTGCTCTCAATGAA----- 994  
Qy 121 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 140  
Db 995 TATCTATTGGACAAGTGCTTACCGTCTTCTCTATTTTATTTGGGACCTTTCAGTATT 1054  
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
Db 1055 GGACATTAGCCCAACATAGAACCTTTGCAATGCAAGAGGGGCGAGCTATGAATC 1114  
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180  
Db 1115 TTCAAGATAATTGATAATGAGCAAGCATTGACAGCTTCTCAACCAAGGGGACACAACCA 1174  
Qy 181 GluSerIleGluThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200  
Db 1175 GACAGTATAATGGAAATTTGGAAATTTAAATAATGTTTACTTCACTACCCATCACGAAGT 1234  
Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
Db 1235 GAAGTTAAGATCTTGAAGGCCCTCAACCTGAAGTGAAGAGCGGCGAGCGTAGCCCTG 1294  
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240  
Db 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240

Db 1295 GTTGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGAC 1354  
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260  
Db 1355 CCCATAGAGGGCGAGGTCACTAGTATCGACGGACAGGACATCAGGACCAATCAATGTGAGTAT 1414  
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 280  
Db 1415 CTGCGGAAATCATTTGGGTGTGAGTCAGSAAACCGTCTGTTCGCCACCAATGTC 1474  
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300  
Db 1475 GAACAACATTCGTATGCGCGAGAAACGTCACCATGATGATAGTAGAAGCTGCAAG 1534  
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
Db 1535 GAAGCCAATGCTATGACTTCATGAAACTGCCCCACAAATTTAACACCTGGTGGT 1594  
Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340  
Db 1595 GAGAGAGGGCGCAGCTGAGTGGGGACAGAAACAGAGATCGCCATTCGCGGGCCCTG 1654  
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
Db 1655 GTCCGCAACCCCAAGATCCTTTTGGATGAGCCACGCTCAGCCTTGGACACAGAAAGC 1714  
Qy 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380  
Db 1715 GAAGCGTGTTCAGCGCTCTGGATAAGGCTAGAGAGGCGGAGCCACCATTTGTGATA 1774  
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
Db 1775 GCTCACCGTGTGTACAGTGGCAATGCTGACGTCACTTGTGGTTTGTGATGTTGTTGTC 1834  
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTrpSerLeu 420  
Db 1835 ATTGTGGACAAAGAAATCATGAAGAGCTCATGAAGAGAGGGCAATTTACTTCAAACT 1894  
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
Db 1895 GTCATGACACAG---ACTAGAGGAATGAATTTGAACAGGAAATTAATGCTTATGAATCC 1951  
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460  
Db 1952 CAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATA 2008  
Qy 461 AspLysAla-----GluGluSerThrGlnSer 469  
Db 2009 AGGAGATCAATTGCGCAGAGATATCCACAGAACACAGACAGAGAGAGACTTAGTTCG 2068  
Qy 470 LysGlu-----IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 486  
Db 2069 AAAGAGATGTGATCAAGATGTCCTATGTTTCTTTTGGCAGATCTCTAAAGCTAAT 2128  
Qy 487 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506  
Db 2129 ATTAGTAATGGCCCTATTATTAGTTGGGTGTACTTTGTGCTGTATAAATGGGTGCATA 2188  
Qy 507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAsp 525  
Db 2189 CAACAGTGTGTGCCATAGTGTTCAAAGATGTAGGGGTTCCTCAAGAGACGACGAC 2248  
Qy 526 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 545  
Db 2249 CATGAACCAACCAACCGAATGTAACTTGTTCCTTCTTCTTCTGTCATGGGAATG 2308  
Qy 546 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 565  
Db 2309 ATTTCTTTTGTACGTACTTCTTCAAGGCTTCACATTTGGCAAGAGCTGGAGAGATCCTC 2368  
Qy 566 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 585  
Db 2369 ACCAAGCGACTCCGATACATGTTCTTCAAAATCCATGCTGCGACAGGATATAAGCTGTTT 2428





```
QY 786 AlaAlaHisLeuPheAlaLeuLeuGluLysProAsnIleAspSerArgSerGlnGlu 805
Db 3029 GCATCCACATCATCAGATCATTTGAGAAATCCCGAGATTGACAGCTACAGACCGAG 3088
QY 806 GlyLysLysProAspThrCysGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
Db 3089 GCGTTGAAGCCTAATGTTAGAGAGAAATGTGAATTTAATGAGTCATGTTCAACTAT 3148
QY 826 ProCysArgProAspValPheIleLeuArgGlyLeuSerIleGluArgGlyLys 845
Db 3149 CCCACCGACCCCAATCCAGTCTTCAGGAGCTGAGCTTCGAGGTGAAGAGGGGCAA 3208
QY 846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 865
Db 3209 ACGCTTCGCTGGTGGGACAGTGGCTCGGGAAGATGACAGTGGTCCAGCTGCTGAG 3268
QY 866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
Db 3269 CGCTTCTACAAACCCCATGGCTGGCAACAGTGTCTTAGATGGCAAGAAATAAACAACTC 3328
QY 886 AsnValGlnTrpLeuArgSerGlnIleAlaTleValProGlnGluProValLeuPheAsn 905
Db 3329 AACGTCCAGTGGCTCCGCGCCACCTGGCAATGTGTCCAGGAGCCCTCTGTTTGAC 3388
QY 906 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 925
Db 3389 TGCAGCATCCAGAGACATCGCTACGGAGACAAACAGCCGTGTCGTCTCATGAGGAG 3448
QY 926 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 945
Db 3449 ATCGTAGGGCGCGCAGGAGGCGCAACATCCACAGTTCATCGACTCACTGCCTGAGAAA 3508
QY 946 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 965
Db 3509 TACAACACCGAGTGGGAGACAAAGGAGCTCAGCTGTGGCGCGGAGACGACGCGCATC 3568
QY 966 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 985
Db 3569 GCCATCGGGCGGCTCGTCAGACAGCTCACATCTTACTTCTGGATGAAGGACATCA 3628
QY 986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
Db 3629 GCTCTGATACGAGAGTGAAGAGTGGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3688
QY 1006 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
Db 3689 CGCACCTGCATTGTGTCGGCACCGGCTGTCCACCATCCAGAACGCGAGACTTGTATGCTG 3748
QY 1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 1045
Db 3749 GTGATTCAAGACGCGCAGGTCAAGGAGCAGCGCACCCACCCAGCAGCTGCTGGCCAGAAA 3808
QY 1046 AspIleTyrPheLysLeuValAsnAla 1054
Db 3809 GGCATCTATTCTCGATGTTCCAGGCT 3835

RESULT 9
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,447
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
; US-08-752-447-1

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3073.50 Matches: 590
Percent Similarity: 74.98% Conservative: 210
Best Local Similarity: 55.30% Mismatches: 224
Query Match: 57.66% Indels: 43
DB: 2 Gaps: 6

US-08-752-409-4 (1-1058) x US-08-752-447-1 (1-4669)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db 1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC 1195
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 1196 TTGGCAGCAATTAGAACTGTGATTGCTATTGGAGGACAAAAGAAAGAACTTGAA----- 1249
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACAAACAAATTTAGAGAGCTAAAGAAATTGGGATAAG 1294
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAGCTATTACAGCCAATTTCTATAGTGTCTCTTCTCGTGTATCATGCTATCTAT 1354
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 123
Db 1355 GCTGTGCGCTTCTGTATGGACACCTTGTCTCTCAGGGAA-----TATTCTATT 1408
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGACAGTACTCACTGTATTCTTTCTGTATTAAATGGGGCTTTTAGTGTGACAGGCA 1468
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
```

```
Db 1469 TCTCCAAGCATTGAACATTTGCAAAATGCAAGAGGACGCTTATGAATCTTCAAGATA 1528
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1529 ATTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCACAAACAGATATATT 1588
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAGTTAAG 1648
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyLeuThrValAlaLeuValGluLeu 223
Db 1649 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGCAGACGGTGGCTGGTTGGAAC 1708
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 1709 AGTGGCTGTGGNAGAGCACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1768
Qy 244 GlyPheIleValMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1769 GGGATGCTCAGTGTTCATGTCAGCAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAA 1828
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1829 ATCAATGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGCCACCAAGATAGCTGAAACATTT 1888
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1889 CGCTATGGCCGGTGAATAATGTCACATGATGATGAGATTTGAGAAAGCTGTCAAGGAAGCAAT 1948
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1949 GCCTATGACCTTTATCATGAACTGCCTCATAAATTTGACACCCCTGTTGGAGAGAGGG 2008
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsn 343
Db 2009 GCCCAGTTGAGTGGTGGCAGAAAGCAGAGATCGCCATTGTCACGCTGCTGGTTCGCAAC 2068
Qy 344 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 2069 CCCAAGATCTCTGCTGATGAGGCCAGCTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 2129 GTTCAGTGGCTCTGGATAAGGCCAGAAAGTTCGACCAACATTGTGATAGTCTCATCGT 2188
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db 2189 TTTGCTACAGTTCGTAATGCTGACGTCATCGCTGTTTCGATGATGAGTCAATTGTGGAG 2248
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db 2249 AAGGAAATCATGATGAATCATGAAGAAAGGATTTACTTCAAACTTGTCAAAATG 2308
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db 2309 CACAGCAGGAAATGAAGTTGAATTAGAAAATGACGCTGATGAATCCAAAAGTGAAATT 2368
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452
Db 2369 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCCATTAAGAAAAAGATCAACT 2428
Qy 453 ---LysSerIleLysSerAspPheIleAspLysAlaGluSerThr---GlnSerLys 470
Db 2429 CGTAGGAGTGTCCGTGGATCACAGGCCCAACACAGAAAGCTTAGTACCAAGAGGCTCTG 2488
Qy 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 490
Db 2489 GATGAAGATATACCTCCAGTTTCTTTGGAGGATTAATGAAGCTAAATTTAACTGAATGG 2548
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db 510
```

```
2549 CCTTATTTTGTGTTGGTGTATTTTGTGCCATATATAATGGAGGCTGCAACCAGCATTT 2608
511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
2609 GCAATAATATTTTCAAGAGATTATAGGGTTTTTACAAGAATTTGATGATCTGAAACAAA 2668
530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
2669 CGACAGAATAGTAACCTGTTTTCACATATTGTTTCTAGCCCTTGGAAATATTATTTCTTT 2728
550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
2729 ACATTTTCTTCAGGGTTTCACATTTGGCAAAAGTGGAGAGATCTCTCAACAGCGGCTC 2788
570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589
2789 CGATACATGTTTTCGGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAA 2848
590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
2849 AACACCACTGGAGCAATGACTACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT 2908
610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
2909 ATAGGTTCCAGGCTGCTGTAATTAACCCAGAAATATAGCAATCTTGGGACAGAAATATT 2968
630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
2969 ATATCTTTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGACCCATCAT 3028
650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
3029 GCAATAGCAGGAGTGTGTAATGAAATGTTTGTGGCAAGCACTGAAAGATAGAAA 3088
670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuLeuAsnIleArgThrIleVal 689
3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCGAAACCGTGT 3148
690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
3149 TCTTTGACTCAGGACAGAAAGTTTGNACATATGATGCTCAGAGTTTGGAGGTACCATAC 3208
710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
3209 AGAAACTCTTTGAGGAAGCACACATCTTTGGAATTTACATTTTCTTCCACCAGGCAATG 3268
730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
3269 ATGTAATTTTCTATGCTGGATGTTTTCGGTTTGGAGCCTACTTGGTGGCACATAAATC 3328
750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
3329 ATGAGCTTTGAGAGTCTCTGTAGTATTTTTCAGCTGTTGTTGTTGTCCTGTCGCGGTG 3388
770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789
3389 GGGCAAGTCAGTTCATTTGCTCTGACTATGCAAAAGCAAAATATATCAGCAGGCCACATC 3448
790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro 809
3449 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGTACAGCAGCGAAGGCCCTTAATGCGC 3508
810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
3509 AACACATTGGAAGAAATGTCACATTTGTTGAGTTGTTTCACTATCCACCCACCG 3568
830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
3569 GACATCCAGTGTCTCAGGACTGAGCCTGGAGGTGAAGAGGCCAGACGCTGGCTCTG 3628
850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
3629 GTGGCAGCAGTGGTGTGGGAAGAGCACAGTGGTCTCAGCTCTCTGGAGCGGTCTTACGAC 3688
```

```
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db 3689 CCCTGGCAGGAAAGTGTCTGTATGTCGCAAGAAATAAGGACCTGAATGTCAGTGG 3748
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db 3749 CTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTGTTTGACTCGACATGCT 3808
QY 910 GluAsnIleAlaTrpGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 929
Db 3809 GAGAACATTGCTATGAGACAACAGCCGGGTGTGTACAGGAGAGATCGTGAGGGCA 3868
QY 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTrpAsnThrGln 949
Db 3869 GCAAAGGAGCCCAACATACATCGCTTCATCGAGTCACGTCTAATAATATAGCACTAAA 3928
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db 3929 GTAGGAGACAAAGGAACTCAGCTCTCTGTGGCCAGAAACAAGCATTTGCCATAGCTGT 3988
QY 970 AlaLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3989 CGCCTTTAGACAGCCCTCATATTTTGTGGATGAAGCCACGTGAGCTCTGGATACA 4048
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 4049 GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAACCCAGAGAGCCGACCTGCATT 4108
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuValValLeuHisAsn 1029
Db 4109 GTGATTCTCACCGCTGTCCACCATCCAGAAATCCAGACTTAATAGTGTGTTCAGAAAT 4168
QY 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTrpPhe 1049
Db 4169 GCGAGTCAAGGAGCATGGCCCATCAGCATCAGCAGCTGTGGCACAGAAAGCACTATT 4228
QY 1050 LysLeuValAsnAlaGlnSer 1056
Db 4229 TCAATGTCAGTGCAGGCT 4249

RESULT 10
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechtner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; PRIOR APPLICATION DATA:
; CLASSIFICATION:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357nan, Kevin E
; REGISTRATION NUMBER: 35,303
```

```
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
; US-09-316-167-1
Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3073.50 Matches: 590
Percent Similarity: 74.98% Conservative: 210
Best Local Similarity: 55.30% Mismatches: 224
Query Match: 57.66% Indels: 43
DB: 4 Gaps: 6
US-09-873-409-4 (1-1058) x US-09-316-167-1 (1-4669)
QY 4 SerLeuThrSerLysGluLeuSerAlaTrpSerLysAlaGlyAlaValaGluVal 23
Db 1136 TCATTTACTGATAAGAACTCTTAGGTATGCAAAAGCTGAGCAGTAGCTGAAGAGGTC 1195
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db 1196 TTGGCAGCAATTAGAACTGTGATTGCTATTGGAGGACAAAAGAAAGAACTTGAA----- 1249
QY 44 PheLeuLeuAsnIleThrArgTrpAlaTrpPheTrpPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
QY 64 ValLeu***PheValArgTrpThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGTACACAAAATTTAGAGAGAGCTAAAGAAATTTGGATAAG 1294
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTrpPhePheMetAsnGlyThrTrp 103
Db 1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGCTATCTAT 1354
QY 104 GlyLeuAlaPheTrpTrpGlyThrSerLeuIleLeuAsnGlyGluProGlyTrpThrIle 123
Db 1355 GCTGTGCGCTTCTGTTATGGGACCACTTGCTCTCTCAGGGGAA-----TATTCTATT 1408
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTrpCysIleGlyAlaAla 143
Db 1409 GGACAAGTACTACTGTATTCTTTCTGTATTAAATTTGGGGCTTTTAGTTGACAGGCA 1468
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 1469 TCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTrpLysProGluSerIle 183
Db 1529 ATTGATAATAAGCCAAGTATTGACGATTTCGAAGAGTGGGCACCAACCAAGATAATT 1588
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTrpProSerArgProSerIleLys 203
Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAAGAGTTAAG 1648
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValaLeuValGlyLeu 223
```



Db 1649 ATCTGAAGGCGCTGAACTCTGAAGGTGCAGAGTGGCAGCGGTGGCCCTGGTTGGAAC 1708  
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGTGCAGAGGCTCTATGACCCACAGAG 1768  
QY 244 GlyPheIleMetValAspGluLeuAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1769 GGGATGCTAGTGTGTATGACACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGAA 1828  
QY 264 HistGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1829 ATCAATGGTGGTGGTGAAGTCAAGAACCTGTATGTTTCCACACGATAGCTGAAACATT 1888  
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1889 CGCTATGGCGGTGAATATGTCACCATGGATGAGATGCAAGAACGTCACAGGAAGCCAA 1948  
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323  
Db 1949 GCCTATGACTTATCATGAACCTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 2008  
QY 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 2009 GCCCAGTTGAGTGGTGGCAGAAGCAGAGATCGCCATTCGACGTGCCCTGGTTGCCAAC 2068  
QY 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 2069 CCCNAGATCTCTGCTGGATGAGGCCACGTCAGCCTTGACACAGNAAGCGAGCAGTG 2128  
QY 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 2129 GTTCAGGTGCTCTGGATAAGGCAGAAAAAGTTCGGACCACTTGTGTAGTCTCATCGT 2188  
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 2189 TTGCTACAGTTCGTAAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATGTTGGAG 2248  
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 2249 AAGGAATCATGATGACTCATGAAGAGAAAGGCAATTTACTTCAAACTGTGCACATG 2308  
QY 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
Db 2309 CAGACACAGGAAATGAAGTTGAATTAGAAATGCAGCTGATGAATCCAAAAGTGAAT 2368  
QY 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 2369 GATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGATCCAAAGAAAAAGATCAACT 2428  
QY 453 ---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470  
Db 2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAGCTTAGTACCAAGAGGCTCTG 2488  
QY 471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 490  
Db 2489 GATCAAGATATACCTCCAGTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548  
QY 491 PropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2549 CCTATTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2608  
QY 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 2609 GCAATAATATTTCAAAGATTATAGGGGTTTTCAGAAATTTGATGATCTCGTGAACAAA 2668  
QY 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2669 CGACAGATAGTAATGTTGTTTCACTATTGTTTACCCCTTGGAAATATTCTTTTATT 2728  
QY 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569

Db 2729 ACATTTTTCTTCAGGGTTTCAATTTGGCAAAAGCTGGAGAGATCTCTACCAAGCGGCTC 2788  
QY 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589  
Db 2789 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTTTCATGACCTAAA 2848  
QY 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2849 AACACCACTGGAGCATTTGACTACAGGCTCGCAATGTAGTCTCAAGTTAAAGGGCT 2908  
QY 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2909 ATAGGTTCCAGGCTCTGTAAATACCAAGATATAGCAAACTCTGGGACAGAAATATT 2968  
QY 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuLeuSerIleAlaProValLeu 649  
Db 2969 ATATCTTTCATCTATGTTGGCAACTAACACTGTTTACTCTTAGCAATTTGACCCATCA 3028  
QY 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 3029 GCAATAGCAGGAGTGTGTTGAAATGAAATGTTTGTCTGGAACAAGCACTGAAAGATAAGAA 3088  
QY 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 3089 GAATAGAGGTGCTGGAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTT 3148  
QY 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 3149 TCTTTGACTCAGGACAGAGTGTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 3208  
QY 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 3209 AGAAACTCTTTGAGGAAGCAGACATCTTTGGAAATTTACATTTTCTTCCACCCAGGCAATG 3268  
QY 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrIleGlnAlaGlyArg 749  
Db 3269 ATGATTTTCTTATCTGATGATGTTTCCGGTTTGGAGCCTTACTTGTGGCACAATACTC 3328  
QY 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 3329 ATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTTCTTGTGCTGCTGCGGTG 3388  
QY 770 GlyThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 3389 GGCAAGTCAGTTCATTTGCTCTGACTATGCCAAAGCCAAATATATCAGCAGCCACATC 3448  
QY 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3449 ATCATGATCATTTGAAAAAACCCCTTTGATTTGACAGCTACAGCACGGAAGGCTTAATG 3508  
QY 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysAlaArgPro 829  
Db 3509 AACACATTTGGAAGAAATGTCATTTTGTGTAAGTTGTTTCAACTATCCACCCGACCG 3568  
QY 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3569 GACATCCAGTGTCTCAGGACTAGCCTGGAGGTGAAGAGGCCAGCAGCTGGCTCTG 3628  
QY 850 ValGlySerSerClyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3629 GTGGCAGCAGTGGCTGTGGGAAGAGCAGACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGAC 3688  
QY 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889  
Db 3689 CCCTTGGCAGGGAAGTGTGCTGTGATGGCAAGAAATTAAGCGACTGTAATGTTCTAGTGG 3748  
QY 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3749 CTCGAGCAGCACCTGGGCTCGTGTCCAGGAGCCCATCTCTGTTTCTGCTGACGACTGCT 3808  
QY 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3809 GAGAACATTTGCCTATGGAGACAAACAGCCGGTGTGTCTCACAGGAAGAGATCGTGGAGGCA 3868



```
QY 930 AlaAsnAlaAlaAenIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
|||
Db 3869 GCAAGAGGCGCACATACATGCTTCATCGAGTCACCTGCTTAATAATATAGCACTAAA 3928
|||
QY 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArg 969
|||
Db 3929 GTAGGAGCAAGGAAGCACTCAGCTCTCTGGTGGCCAGAAACAAACGCAATGCCATAGCTCGT 3988
|||
QY 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
|||
Db 3989 CGCCTGTGTAGACAGCCCTCATATTTGCTTTGGATGAAGCCAGCTCAGCTCTGGATACA 4048
|||
QY 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
|||
Db 4049 GAAAGTGAAGAGTTGTCAGAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCAAT 4108
|||
QY 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
|||
Db 4109 GTGATTGCTACCGCCCTGTCACCATCCAGAAATCAGACTTTAATAGTGGTGTTCAGAAAT 4168
|||
QY 1030 GlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
|||
Db 4169 GGCAGAGTCAAGAGGATGCGCACGCATCAGCAGCTGCTGGCAGAAAGGCATCTATTTT 4228
|||
QY 1050 LysLeuValAsnAlaGlnSer 1056
|||
Db 4229 TCAATGGTCAGTGTCCAGGCT 4249
|||

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
US-08-583-276-18
```

## Alignment Scores:

```
Pred. No.: 0 Length: 4669
Score: 3056.50 Matches: 588
Percent Similarity: 74.86% Conservative: 210
Best Local Similarity: 55.16% Mismatches: 225
Query Match: 57.35% Indels: 43
DB: 2 Gaps: 6

US-09-873-409-4 (1-1058) x US-08-583-276-18 (1-4669)

QY 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db 1136 TCATTTACTGATAAAGAACTCTTAGCGGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195
|||
QY 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
|||
Db 1196 TTGGCGCAATTAAGAATGTGATTCATTTGGAGGACAAAGAAAGAACTTGA 1249
|||
QY 44 PheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 1249 ----- 1249
|||
QY 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 1250 -----AGGTACAAACAAAAATTTAGAAAGAGCTAAAGAAATGGGATAAG 1294
|||
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 1295 AAAGCTATTACGCCAATATTCTATAGGTGCTGCTTCTCGTGTATCTATGATCTAT 1354
|||
QY 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 123
Db 1355 GCTCTGGCCTTCTGTATGGGACCACCTTGGTCTCTCAGGGGAA-----TATTCTATT 1408
|||
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 1409 GGAAGAAGTACTCCTCTGATTTCTTCTGTATTAATTTGGGCTTTTAGTGTGGCAGGCA 1468
|||
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 163
Db 1469 TCTCCAGCAATTGAAGCAATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
|||
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 1529 ATTGATAATAAGCCAAAGTATTGACAGCTATTCGAGAGTGGGCACAAACACAGATAATATT 1588
|||
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATTCGAAAAGAGTTAAG 1648
|||
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 1649 ATCTTGAGGGCCTGAACCTGAAGGTGCAGAGTGGCAGACGGTCCCTCGTTGGTGAAC 1708
|||
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuArgLeuTyrAspProAspAsp 243
Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCACCTGATGCAGAGGCTCTATGACCCACAGAG 1768
|||
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 1769 GGGATGTCAGTGTTCATGGACACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAA 1828
|||
QY 264 HistGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db 1829 ATCATTTGGTGTGTAGTCAGGAACCTGTATTGTTTCCACACGATAGCTGAACAACATT 1888
|||
QY 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 1889 CGCTATGGCCGTGAATGTCACCATGATGAGATGAGAAAGCTGTCAAGGAAGCCAAAT 1948
|||
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 1949 GCCTATGACTTTTATCATGAAACTGCCTCTATAAAATTTGACACCTCGTTGGAGAGAGAGGG 2008
|||
```



db 4169 GGCAGAGTCAAGCAGCATGGCAGCGCATCAGCAGCTGCTGGCAGAGAAAGGCATCTATTTT 4228

Qy 1050 LysLeuValAsnAlaGln 1055  
          :::||||::: |||  
Db 4229 TCAATGGTCAGTGTCCAG 4246

Db 4229 TCAATGGTCAGTGTCCAG 4246

## RESULT 12

US-08-461-823-1  
Sequence 1 Application US/08461823  
Patent No. 5593840  
GENERAL INFORMATION:  
APPLICANT: Bhatnagar, Satish K.  
APPLICANT: George Jr., Albert L.  
APPLICANT: Nazarenko, Irina  
TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES

Alignment Scores:					
Pred. No.:	3.5e-246	Length:	2726		
Score:	2280.00	Matches:	435		
Percent Similarity:	76.55%	Conservative:	159		
Best Local Similarity:	56.06%	Mismatches:	168		
Query Match:	42.78%	Indels:	14		
DR:	1	Gaps:	4		

US-09-873-409-4 (1-1058) x US-08-461-823-1 (1-2726)

Qy	295	MetGluArgAlaAlaAraGluAlaAsnAlaTyraSpPheIleMetGluPheProAsnLys	314
		.....	
		.....	
Db	2	ATTGGAAGCTGTCAAGGAGCCAAATGCCTATGACTTTATCATGAAACTGCCTATAA	61
		.....	
Qy	315	PheAsnThrLeuValGlyGlyGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle	334
		.....	
		.....	

62 TTTGACACCTTGGTGAGAGAGAGGGCCCACTGTAGTCGGCGCAGAACGAGGATC 121

335 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 354  
::: :::::::::::::::::::::  
122 GCCATTGCACGTGCCCTGGTTCGAAACCCCAAGATCCTCTGCTGGATGAGGCACGCTCA 181

355 AlaLeuAspSerGluSerLysSerAlaValClnAlaAlaLeuGluLysAlaSerLysGly 374  
||||| :::::::::::::::::::::  
182 GCGTTGGACACAGAAGCGAAGTAGTGTTACAGTGGCTCTGGATAGGCCAGAAANGT 241

375 ArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuVal 394  
||||| :::::::::::::::::::::  
242 CGACACCACCATTTGATAGCTCATCTGTTGTCTACAGTTCGTAATGCTGACGTCATCGCT 301

395 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 414  
||||| :::::::::::::::::::::  
302 GGTTTTCGATGATGGAGTCATTTGTGGAGAAAGAAATCATGATGAATCATGAAGAAGAAA 361

415 GlyLeuTyrrSerIleValMetSerCln-----AspIleLysLys 428  
||||| :::::::::::::::::::::  
362 GGCATTACTTCAAACCTTGTCCAAATGCAGACACAGCGAAATGAAGTTGAATAGAAAAAT 421

429 AlaAspGluGlnMetGluSerMetThrThyrrSerThrGluArgLysThrAsnSerLeuPro 448  
||||| :::::::::::::::::::::  
422 GCAGCTGATGAATCCAAAGATGAATTTGATGCCCTGGAAATGCTTCAATGATTCAAGA 481

449 LeuHisSerVal-----LysSerIleLys---SerAspPheIleAsp 461  
::: :::::::::::::::::::::  
482 TCACGTCTAATAAGAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAAGCCCAGAC 541

462 LysAlaGluGlnSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLys 481  
::: :::::::::::::::::::::  
542 AGAAAGCTTAGTACCACAGAGGCTCTGGATCAAAAGTAGTATCCTCCAGTTTCTCTTTGGAGG 601

482 IleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerVal 501  
||||| :::::::::::::::::::::  
602 ATTATGAAGCTAAATTTAACTGAATGGGCTTATTGTTGTGTGTGATTTTGTGGCATT 661

502 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 521  
::: :::::::::::::::::::::  
662 ATAAATGGAGGCTGCACACGACATTTGCATATAATATTTCAAGATTATAGGGTTTTT 721

522 GlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540  
||||| :::::::::::::::::::::  
722 ACAAGAATTGATGATCCTGAAACAAACAGCAGACAATAGTAACTTGTTTTTCACATTATGTTT 781

541 ValIleLeuGlyValIleCysPheValSerThyrrPheMetGlnGlyLeuPheThyrrGlyArg 560  
::: :::::::::::::::::::::  
782 CTAGCCCTTGGAAATTTCTTTTATATACATTTTTTCTCTCAGGGTTTCACATTTGGCNA 841

561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580  
||||| :::::::::::::::::::::  
842 CTTGGAGAGATCTCTACCAAGCGGCTCGATACATGTTTCCGATCCATGCTCAGACAG 901

581 AsplleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600  
||||| :::::::::::::::::::::  
902 GATGTGAGTTGGTTTGATGACCTAAACACCACTGGAGCATTTGACTACCGAGCTCGCC 961

601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620  
||| :::::::::::::::::::::  
962 AATGATCTCTCAAGTTTAAAGGGGCTAGAGTTCCAGGCTTGCTGTAATTACCAGAAAT 1021

621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640  
||||| :::::::::::::::::::::  
1022 ATAGCAATCTTGGGACAGGAATAATTATCTTCTATCTATGTTGGCACTAACACTG 1081

641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660  
||||| :::::::::::::::::::::  
1082 TTACTCTTAGCAANTGTACCCATCATTTGCATATAGCAGGAGTTGTTGAAATGAAATGTTG 1141

661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680  
::: :::::::::::::::::::::  
1142 TCTGGACAAGCATCTGAAAGATAGAAAGAACTAGAAAGTGTCTGGGAAGATCGCTACTGAA 1201

Qy	681	AlaLeuGlnAsnIleA	gThrlleValSerLeuThrArgGluLysAlaPheGlnMet	700
Db	1202	GCAATAGAAACTT	CGCAACCGTGTCTTTGACTCAGGAGCAGAAGTTTGAACATATG	1261
Qy	701	TyrGluClnMetLeu	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleGly	720
Db	1262	TATGCTCAGAGTT	TGCAGGTACATACAGAACTCTTTGAGGAAGACACATCTTTGGG	1321
Qy	721	SerCysTyrAlaPhe	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe	740
Db	1322	ATTACATTTTCC	TTCACCCAGGCAATGATATTTTCTATGCTGGATGTTTCCGGTTT	1381
Qy	741	GlyAlaTyrIleuIle	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr	760
Db	1382	GGAGCCTTACT	TGTGTGCACATAACTCATGAGCTTGTGAGGATGTTCTGTATGATATTTCA	1441
Qy	761	AlaIleAlaTyrGly	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer	780
Db	1442	GCTGTGTC	TCTTGGTCATGCGCGTGGGGCAAGTCAGTTCACTTGTCTCTGCATATGCC	1501
Qy	781	LysAlaLysSerGly	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp	800
Db	1502	AAAGCCAAAATAT	CAGCAGCCACATCATCATGATCATTTGAAAAACCCCTTTCATTGAC	1561
Qy	801	SerArgSerGlnGlu	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu	820
Db	1562	AGCTACAGCAG	GGAAGCCCTTAATGCCGAAACACATTTGGAAGGAATGTCCACATTTGGTGA	1621
Qy	821	ValSerPheTyrPro	CysArgProAspValPheIleLeuArgGlyLeuSerLeuSer	840
Db	1622	GTGTGATTCAACT	TATCCACCCGACCGACATCCACGAGTCTTCAGGAGCTGAGGCTGGAG	1681
Qy	841	IleGluArgGlyLys	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer	860
Db	1682	GTGAAGAAGGCC	CAGACGCTGGCTCTGTGTGGCAGCAGTGGCTGTGTGGGAAGACACAGT	1741
Qy	861	ValGlnLeuLeuGln	ArgLeuTyrAppProValGlnGlyGlnValLeuPheAspGlyVal	880
Db	1742	GTCAGCTCTCTG	AGCGGGTCTTACGACCCCTTTGGCAGGGAAGTGTCTGTGATGGCAAA	1801
Qy	881	AspAlaLysGluLeu	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu	900
Db	1802	GAATAAAGCGAT	GTGATGTTCTAGTGGCTCCGAGCACACCTGGGGCATCTGTCTCCAGGAG	1861
Qy	901	ProValLeuPheAsn	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal	920
Db	1862	CCCATCTGTTT	GACTGCAGCATTTGCTGAGAACATTTGCCATGTGAGACAAACAGCGGGTG	1921
Qy	921	ValProLeuAspGlu	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu	940
Db	1922	GTGTTCAGGAAG	AGAGATCGTGGGGCAGCAAGAGGGCCAAACATACATGTCCTTCATCGAG	1981
Qy	941	GlyLeuProGluLys	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly	960
Db	1982	TCACTGCGCTA	TAAATATAGACTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGTGGTGC	2041
Qy	961	GlnLysGlnArgLeu	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu	980
Db	2042	CAGAAAACAGCG	ATGCGCATGCTCGTGCCCTTTGTATAGACAGCCTCATATTTGCTGTTTG	2101
Qy	981	AspGluAlaThrSer	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp	1000
Db	2102	GATGAGCCACG	TACGCTCTGGGATCAGAAAGTGAAGGTTGTCCAAAGACCCCTGGAC	2161
Qy	1001	LysAlaAlaArgThr	GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn	1020
Db	2162	AAAGCCAGAGA	GCGCCACTCTCATTTGTGCTTCACCGCTGTCTCCACCATCCGAAT	2221
Qy	1021	AlaAspLeuIleVal	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu	1040
Db	2222	GCACACTTAAT	ATAGTGTGTTTCAGATGCGCAGAGTCAAGAGCATGGCAGCATCAGCAG	2281

```

Qy 1041 leuLeuA9rGAsnArGaSpIleTyPheLySeuValAsnAlaGlnSer 105
||||| :: ||||| ||||| :|||::| |||||
Db 2282 CTGCTGGCAGAAAGGCATCTATTTTTCAATGTGTCAGTGCCAGGCT 2329

RESULT 13
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: gDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
; US-08-996-545-1

Alignment Scores:
Pred. No.: 6,76e-222 Length: 4002
Score: 2067.00 Matches: 447
Percent Similarity: 59.43% Conservative: 199
Best Local Similarity: 41.12% Mismatches: 381
Query Match: 38.78% Indels: 60
DB: 2 Gaps: 9

US-09-873-409-4 (1-1058) x US-08-996-545-1 (1-4002)

Qy 2 ValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAlaGlyAlava
:||||| :: ||||| ||||| :|||::| |||||
Db 817 ATCATCAAGTACAGCAAAAAGTCGCTTTGACAGCTACGCGTCAGCGCGCACTGT
:||||| :: ||||| ||||| :|||::| |||||
Qy 22 GluValLeuSerSeriIleArgThrValIleAlaPheArgAlaGlnGluLysGl
:||||| :: ||||| ||||| :|||::| |||||
Db 877 GAGGTTCATCAGCTCCATCAGAAATGCCACAGCGTTTGGCACCCCAAGACAGCT
:||||| :: ||||| ||||| :|||::| |||||
Qy 42 ArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyrPheProGlnTr

```

Db 936 ----- 936  
QY 62 SerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 81  
Db 937 -----CAGTATGAGGTCCACTTAGACGAGCTGAGAAATGGGGA 975  
QY 82 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 101  
Db 976 ACAAGAACCAAGATTGTCATGGGTTTCATGATTGGCCCATGTTTGCCCTTATGCTACTCG 1035  
QY 102 ThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 121  
Db 1036 AACTACGCTCTTGGCTTTCTGATGGGTTCTGCTTTCTCTGCTAGTAGTGT-----GCATGC 1089  
QY 122 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 141  
Db 1090 GATGTGGGTGATATCTCACAGTTCTCATGGCCATCTTGATCGGATCGTCTCTCTCTGGG 1149  
QY 142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161  
Db 1150 AACCTTAGTCCAAATGCTCAAGCATTTACAAACGCTGTGGCGCGCGCAAGATATTT 1209  
QY 162 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181  
Db 1210 GGAACGATCGATCGACCGACGTCCTCCCATTTAGATCCATATTGCAACGAAGGGAAGACGCTCGAC 1269  
QY 182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201  
Db 1270 CATTTTAGGGCCACATTGAGTTAGCAATGTCAAGCATATTTACCATCTAGACCCGAG 1329  
QY 202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221  
Db 1330 GTCACCGTCATGGAGGATGTTCTCTGTCATGCGCGCTGGAAAAACAACCGCTTTAGTC 1389  
QY 222 GlyLeuAsnGlySerGlySerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241  
Db 1390 GGCCCTCTGGCTCTGGAAAAAGTACGGTGGCTGGCTTGGTGGAGCATTTCTACATCGCT 1449  
QY 242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 261  
Db 1450 GTTCGGGTACGGTTTGTCTGGATGGCCATGACATCAAGGACCTCAATCTCCGTGGCTT 1509  
QY 262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 281  
Db 1510 CGCCACACAGATCTCTTGGTTAGCCAGGAGCTGTCTTTTGGCAGCAGCATTTATAAG 1569  
QY 282 AsnIleLysTyrGly-----ArgAspAspValThrAsp 292  
Db 1570 AATATTAGGCACGGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGTCCGG 1629  
QY 293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312  
Db 1630 GAATCTCATCGAAGACGGGCAAAATGGCGAATGCTCATGACTTTATTTACTGCTTCGCT 1689  
QY 313 AsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332  
Db 1690 GAAGGTATGAGACCAATGTTGGCAGCGTGGCTTTCTCTCTTCAGGTGGCCAGAAACAG 1749  
QY 333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352  
Db 1750 CGCATTTGCAATCGCGCGTGGCTGTTAGTACCCCAAAATTCCTGCTCTCTGATGAAGCT 1809  
QY 353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372  
Db 1810 ACTTCGGCTTGGACACAAATCCGAAGCGGTGGTTCAAGCAGCTTTGGAGAGGGCAGCT 1869  
QY 373 LysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392  
Db 1870 GAAGGCCGAATCTATTGTGATCGCTCATCGCTTCCACGATCAAAACGGCGCACAAAC 1929  
QY 393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412  
Db 1930 ATTCGTGGTCTGGTCAATGGCAAAATGCTGNAACAAGAACTCACGATGAATTTGTTGAC 1989

QY 413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432  
Db 1990 CGCGGAGCGCTTATCGCAAACTTTGGAGGCTCAACGTATCAATGAACAGAGGAAGCT 2049  
QY 433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHis 450  
Db 2050 GAGCGCTTGAGGACGCCGACGCTGAGGATCTCAGAAATGCAGATATTGCCAAATCAAA 2109  
QY 451 SerValLysSerIleLysSerAspPhe-----IleAspLys----- 462  
Db 2110 ACTGCGTCAAGCGCATCTCCGATCTCGACGGAAAAACCAACCATTTGACCGACGGCG 2169  
QY 463 AlaGluGluSerThrGlnSerLysGluIleSer-----LeuProGlu 476  
Db 2170 ACCCAAGTCTGTTCCAGCGGATCTTCTTAAAGACCCCCCGAACAACATCCGAAA 2229  
QY 477 ValSerLeuLysIleLeuLys-----LeuAsnLysProGluTyrProPhe 492  
Db 2230 TACTCATTTAGGACGCTGCTCAAAATTTGTGCTTCTCAACGCGCTGAAATCCCGTAC 2289  
QY 493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512  
Db 2290 ATGCTCATCGCTCTTGTCTTCTCAGTTAGTGGTGGGCCAACCCACGCAAGCAGTG 2349  
QY 513 IlePheAlaLysIleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLys 530  
Db 2350 CTATATGCTAAACCCATCAGCACACTCTCGTCCAGAAATCACAATATAGCAAGCTCGA 2409  
QY 531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550  
Db 2410 CATGATCGGATTTCTGGTCAATGATGTTCTTCTGTTGTTGATTCATTCAGTTTATCAG 2469  
QY 551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570  
Db 2470 CAGTCAACCAATGGTGTGCTGATTTGGCGTATGCTCCGAGAGACTTATTCCTCGCGGAGA 2529  
QY 571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsn 590  
Db 2530 AGCACTGCCTTCGGACGATACTCCGTCAGACATGCTTCTTTCACAGGAAGAGAAT 2589  
QY 591 SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610  
Db 2590 AGACCGCGCTCTGACCTCTTCTGTCGACGAGCAAGCATCTCTCCGTTGTAGC 2649  
QY 611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630  
Db 2650 GGTGTGACTTAGGCAAGTCTTGTGATGACCTCCACGACCTTAGGAGCGGTATCATATT 2709  
QY 631 SerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650  
Db 2710 GCCTGGCGATTGGTGGAAATTTGGCTTAGTTGTTATCTCGTTGTCGGTCTCTCTG 2769  
QY 651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670  
Db 2770 GCATGGGTTTCTACCGATTTCTATATGATAGCCAGTTTCAATCACGCTCAAGCTTGT 2829  
QY 671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSer 690  
Db 2830 TATGAGGATCTCAAACTTGTCTCGAGGCTACATCGTCTATCCGACAGTTGGTGCA 2889  
QY 691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArg 710  
Db 2890 TTAACCGGGAAGGAGTGTCTGGAGATTTACATGCCCGCTTGCACCAAGGAGG 2949  
QY 711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730  
Db 2950 ACCAGCTAATCTCTGTCTTGGGTCAATCCCTGTATATATCGTCTCGTCCGACGAGCTTGT 3009  
QY 731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMet 750  
Db 3010 TTCTTCTGCTGCTCGGTTTGGTACGGAGGACACTTCTTGTGTCCACGAGATAT 3069





Db 1090 GAUGUGGGGAUUAUUCACAGUUCUUCAGUGGCAUCUUGAGUGGAUUCUUCUGGGG 1149  
Qy 142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161  
Db 1150 AACGUUAGUCCAAUUGCUAAGCAUUAUCAAACGCUGUGCGCGCCGCAAGAAUUAU 1209  
Qy 162 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181  
Db 1210 GGAACGAUCCAGCGCAGUCCCAUUAUAGUCCAUUAUUGAAACGAAGGAGAGCGUCGAC 1269  
Qy 182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201  
Db 1270 CAUUUUGAGGGCCCAUUGAGUACGCAUGUCAAAGCAUUAUUAUCCCAUAGACCGGAG 1329  
Qy 202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221  
Db 1330 GUACCGGUCAGGAGGAUUGUUCUGUCAAUGCCCGCGGAAACCAACCGCUUAGUC 1389  
Qy 222 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241  
Db 1390 GGCCCUUCUGGUCUGGAAAAAGUACGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 1449  
Qy 242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaIleAsnValArgHisTyr 261  
Db 1450 GUUCGGCGUACGGUUGUGGAGUAGGCGGCAUAGCAUAGCAUCCUACUCCGUGGCUU 1509  
Qy 262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSerAsn 281  
Db 1510 CGCCACAGAUUCUUGUUGUAGCCAGGAGGUGUUCUUGGCGACGAGAUUAUAG 1569  
Qy 282 AsnIleLysTyrGly-----ArgAspAspValThrAsp 292  
Db 1570 AAUAUAGGCGACGGUCUUCAGCGCACAAAGUACGAGAAUAGUCCGAGGAUAGGUCGG 1629  
Qy 293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312  
Db 1630 GAACUACAGAGAACCGCGGCAAAUAGCGGAAUAGCGCAUUAUUAUUCUGGCUUCGCU 1689  
Qy 313 AsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332  
Db 1690 GAAGUUAUGAGACAAUUGUUGGACGCGGUGUUCUUCUUCAGUGGCGCAGAAACAG 1749  
Qy 333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352  
Db 1750 CGCAUUGCAAUCCGCGCGGUGUUGUAGUAGACCCAAUUAUUCUGUCCUGGAGUAGAC 1809  
Qy 353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372  
Db 1810 ACUUCGCGCUUGGACACAAAUCCGAAUCCGAGCGGUGUUAAGCAGCUUUGGAGGCGCAG 1869  
Qy 373 LysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392  
Db 1870 GAAGCGCAACUACUUAUGUAGUGCUCAUCGCGCUUUCAGCAUCAAACCGCGCACAC 1929  
Qy 393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412  
Db 1930 AUUGUGGUUCUGGCAUUGGCAAAUUAUUGCUGAACCAAGCAUCCACGAUAAUUGUUGAC 1989  
Qy 413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432  
Db 1990 CGCGGAGCGCUUAUUGCAAAUUCUUGGAGGCGUCAAAGCUUAUUAUAGAAAGAGAGCU 2049  
Qy 433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHis 450  
Db 2050 GACGCUUGGAGGACGCGGAGUAGUAGUACUAGUAGUAGUAGUAGUAGUAGUAGUAGU 2109  
Qy 451 SerValLysSerIleLysSerAspPhe-----IleAspLys----- 462  
Db 2110 ACUGCGUACAGCGCAUCCAGUUCGACGCGAAACCCCAACCAUAGCCGACCGCGG 2169  
Qy 463 AlaGluGluSerThrGlnSerLysGluIleSer-----LeuProGlu 476  
Db 2170 ACCCACAGUCUGUUCAGCGGCAUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2229

Qy 477 ValSerLeuLeuLysIleLeuLys-----LeuAsnLysProGluTyrProPhe 492  
Db 2230 UACUCAUAUUGAGCGCUGCUCAAAUAUUGUUGUUCUUAACCGCCUGAAAUCGCGUAC 2289  
Qy 493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512  
Db 2290 AUGCUCAUGCGGUGUGUUCUUCAGUUGUAGUGUGGUGGCGCAACCCACGACGAGUG 2349  
Qy 513 IlePheAlaLysIleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLys 530  
Db 2350 CUUAUGCUUAAGCCCAUAGCAGACACUCUCGCGCCAGAAUCAAUAUAGCAAGCUUCGA 2409  
Qy 531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550  
Db 2410 CAUGAUGCGGAUUCUGGCAUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2469  
Qy 551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570  
Db 2470 CAGUCAACCAUUGGUGCUGCAUUGGCGUAGUCCGAGAGACUUAUUCGUCGCGGAGA 2529  
Qy 571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsn 590  
Db 2530 AGCACUGCCUUGCGGAGCAUACUCCGCAAGCAUUGCUUUAUUGCAAGGAGAGAAU 2589  
Qy 591 SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610  
Db 2590 AGCACCGCGCGCUGACCUUUCUGUCCACCGAGCAAGCAUCCGCGGUGUAGC 2649  
Qy 611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630  
Db 2650 GGUGUACUCUAGGCGACGAUUGUAGUCCACACCCUAGGCGGCUAUAUUAU 2709  
Qy 631 SerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650  
Db 2710 GCCUGCGCAUUGGUGGAAUUGGCUUUAUUGUUGUUGUUGUUGUUGUUGUUGUUG 2769  
Qy 651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670  
Db 2770 GCAUGCGGUUUCUACGAUUCUAUUGUAGUCCGCAUUAUUAUUAUUAUUAUUAU 2829  
Qy 671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuLysAsnIleArgThrIleValSer 690  
Db 2830 UAUGAGGGAUUGCAAAUUCUUGCGAGGCUAUAUUGUUAUUGUUAUUGGCUAUAUUA 2889  
Qy 691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArg 710  
Db 2890 UUAACCGGAAAGGAGUUGUUGGAGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2949  
Qy 711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730  
Db 2950 ACCAGUCAAUUCUUGUUGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3009  
Qy 731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleIleGlnAlaGlyArgMet 750  
Db 3010 UUCUUCUGUUGCGUCGCGGUAUUGUUAUUGGAGGAGACUUAUUGGUGUACCAAGAGUA 3069  
Qy 751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770  
Db 3070 GACAUAUUCGCUUUAUUGUUAUUGUUAUUGUUAUUGUUAUUGUUAUUGUUAUUG 3129  
Qy 771 LysThrIleValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790  
Db 3130 ACCGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3189  
Qy 791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810  
Db 3190 CGACUGUUCACGAAAGCCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3249  
Qy 811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830  
Db 3250 ACGGUGAGGAGGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3309





Qy 182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201  
Db 1270 CATTTTGGAGCCACATAGTACGCAATGTCAGCAATATTTTACCCATCTTAGACCCGAG 1329  
Qy 202 IleLysIleLeuLysGlyLysAsnArgIleLysSerGlyGluThrValAlaLeuVal 221  
Db 1330 GTCCACGTCATGGAGGATGTTCTGTCAATGCGCGCTGGAAAAACAACCGCTTTAGTC 1389  
Qy 222 GlyLeuAsnGlySerGlyLysSerThrValValGluLeuLeuGlnArgLeuTyrAspPro 241  
Db 1390 GCGCCCTCTGGCTCTGGAAAAAGTACGGTGTGCGCTTGGTTGAGCGATCTACATCGCT 1449  
Qy 242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 261  
Db 1450 GTTCGCGGTACGGTGTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTT 1509  
Qy 262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 281  
Db 1510 CGCCAAACAGATCTCTTTGGTTAGCCAGAGCGCTGTTCTTTTGGCAGCAGCATTTATAAG 1569  
Qy 282 AsnIleLysTyrGly-----ArgAspAspValThrAsp 292  
Db 1570 AATATTAGCGGTCTCATCGGCACAAAGTACAGAAATGAATCCGAGGATAAGGTCGG 1629  
Qy 293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312  
Db 1630 GAACCTCATGAGAACGGCGGCAAAATGCGGAATGCTCATGACTTTATTACTCGCTTCGCT 1689  
Qy 313 AsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332  
Db 1690 GAAGGTTATGAGACCAATGTTGGGCGAGCGGTGCTTCTCTTTCAGGTGGCCAGAAACAG 1749  
Qy 333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352  
Db 1750 CGCAATGCAATCGCGGTGCGTGTAGTACCCCAAAATCTGCTCTCGATGAAGCT 1809  
Qy 353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372  
Db 1810 ACTTCGGCTTTGGACACAAATCCGAAGCGGTGTTCAAGCAGCTTTGGAGAGGCGAGCT 1869  
Qy 373 LysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392  
Db 1870 GAAGCGCGAATCTATTGTGATCGCTCATCGCTTTCCAGATCAAAACGGCGCACAAAC 1929  
Qy 393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412  
Db 1930 ATTGTGTTCTGGTCAATGGCAAAATGCTGAACAGAACTCACGATGAATGTTGATC 1989  
Qy 413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432  
Db 1990 CGCGGAGCGCTTATCGCAAACTTTGTGAGGCTCAACGATATCAATGAACAGAGAGCT 2049  
Qy 433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHis 450  
Db 2050 GACGCTTTGGAGAGCGCGAGCTGAGGATCTCAGAAATGAGATATGCGCAAAATCAAA 2109  
Qy 451 SerValLysSerIleLysSerAspPhe-----IleAspLys----- 462  
Db 2110 ACTCGCTACGCGCATCATCGATCTCGACGGAAAAACCCACAACTTACCGCGCGGGC 2169  
Qy 463 AlaGluGluSerThrGlnSerLysGluIleSer-----LeuProGlu 476  
Db 2170 ACCCACAAGTCTGTTTCAGCGCGATTTCTTTAAAGACCCCGGAAAACTCCGAAA 2229  
Qy 477 ValSerLeuLeuLysIleLeuLys-----LeuAsnLysProGluTyrProPhe 492  
Db 2230 TACTCATATGAGCGCTGCTCAAAATTTGTTGCTTCTTCAACCGCCCTGAAATCCCGTAC 2289  
Qy 493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512  
Db 2290 ATGCTCATCGCTCTGTTCTTCTAGTGTAGCTGGTGGTGGCCAAACCCAGCAAGCGATG 2349  
Qy 513 IlePheAlaLysIleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLys 530

Db 2350 CTATATGCTAAAGCCATCAGCACACTCTCGCTCCCGAATCACAATATATGCAAGCTTCGA 2409  
Qy 531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550  
Db 2410 CATGATCGGATTTCTGTCATTTGTTCTTCTGTTGTTGTTATCATTCAGTTTATCAGC 2469  
Qy 551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570  
Db 2470 CAGTCAACCAATGGTGTGCTGCTATTTGCGTATGCTCCGAGAGACTTATTCGTCGCGCAGA 2529  
Qy 571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsn 590  
Db 2530 AGCACTGCTCTTCGGACGATACCTCCGTCACAGACATTTCTTCTTGACAGGAGAGAAAT 2589  
Qy 591 SerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610  
Db 2590 AGACCGCGCGCTCTGACCTCTTTCTGTCACCGAGCAAGCAATCTCTCCGTTGTTAGC 2649  
Qy 611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630  
Db 2650 GGTGTGACTCTAGCAGCATCTTGATGACCTCCACGACCTAGGAGCGGTATCATTTATT 2709  
Qy 631 SerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650  
Db 2710 GCCCTGGCGATTGGGTGGAATTTAGTTTGTATCTCGTTGTCGGTTCCTCTCTCTG 2769  
Qy 651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670  
Db 2770 GCATGCGGTTTTCACGATCTATGCTAGCCAGTTTCAATCAGCTCCAGCTTGTCT 2829  
Qy 671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSer 690  
Db 2830 TATGAGGATCTGCACAACTTTGCTTCGAGGCTACATCGTCTATCCGACAGTTGCGTCA 2889  
Qy 691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArg 710  
Db 2890 TTAAACCGGAAAGGATGTCTGGAGATTTACATGCCAGCTTGACGCACAAAGGCGAG 2949  
Qy 711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730  
Db 2950 ACAGCTAATCTCTGCTTGGAGGTATCCCTGTTATATGCTGCTCGCAGGACCTTGT 3009  
Qy 731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArgMet 750  
Db 3010 TTCTTCTGCGTTCGCTCGGTTTGTACGAGGAGACACTTCTTGTGTACCACGAGTAT 3069  
Qy 751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770  
Db 3070 GACATTTTCGCTTCTTGTGTTCTCTCGAGATTTCTTGTGTTGCTCAATCGCGGGC 3129  
Qy 771 LysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790  
Db 3130 ACCGCTCTTTCTTTCGACAGACATGGGCAAGCGGCAAGAAATCGCGCCCGCAATTCGA 3189  
Qy 791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810  
Db 3190 CGACTGTTCCACGAAAGCCCAAAATGATGATGTTGTTCTGAAGAGGCGGAGAGCTCGAA 3249  
Qy 811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830  
Db 3250 ACGGTGGAAGGTGAAATTCGAATTTAGGAACGTGCACCTTCAGATACCCGACCCGCCAGAA 3309  
Qy 831 ValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheVal 850  
Db 3310 CAGCCTGCTCTGCGCGCTTGGACCTGACCGTGAAGCGCTCGAACAATATGTTGCGCTTGT 3369  
Qy 851 GlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspPro 870  
Db 3370 GGACCCAGCGGTGTTGGCAAGAGTACCACCATTTGCTTGTGAGCGCTTTTACGATCGC 3429  
Qy 871 ValGlnGlyGlnValLeuPheAspGlyValAlaSerAlaLysGluLeuAsnValGlnTrpLeu 890



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 385.104 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-4

Perfect score: 5330

Sequence: 1 WVISUTSKLSAYSAGAVA.....QELLNRDIYFKLVNAQSVQ 1058

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115424\_19290/app\_query.fasta\_1.7544  
-DB=Published Applications NA -QFIX=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat\_27032003\_115424\_19290  
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5328	100.0	3177	10	US-09-873-409-12
2	5328	100.0	3702	10	US-09-873-409-13
3	5159.5	96.8	3621	10	US-09-873-409-14
4	4561.5	85.6	2856	10	US-09-873-409-10

5	3456	64.8	2066	10	US-09-873-409-9	Sequence 9, Appli
6	3111.5	58.4	4317	9	US-10-044-671-1	Sequence 1, Appli
7	3106.5	58.3	4369	10	US-09-769-097-1	Sequence 1, Appli
8	3106.5	58.2	4425	10	US-09-769-097-3	Sequence 3, Appli
9	3100.5	58.2	3860	10	US-09-866-866A-1	Sequence 1, Appli
10	3100.5	58.2	3860	10	US-09-866-866A-3	Sequence 3, Appli
11	3100.5	58.2	8630	10	US-09-306-417-1	Sequence 1, Appli
12	3100.5	58.2	8630	10	US-09-306-417-2	Sequence 2, Appli
13	3090.5	58.0	4788	10	US-09-866-866A-7	Sequence 7, Appli
14	3085	57.9	3924	10	US-09-880-107-2299	Sequence 2299, Ap
15	3085	57.9	4189	10	US-09-866-866A-5	Sequence 5, Appli
16	3084	57.9	4643	9	US-10-072-621-2	Sequence 2, Appli
17	3044.5	57.1	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
18	3043	57.1	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	2488.5	46.7	5010	10	US-09-917-800A-483	Sequence 483, App
20	2103	39.5	4175	10	US-09-749-340-3	Sequence 3, Appli
21	2072.5	38.9	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
22	2067	38.8	4002	9	US-09-758-828-1	Sequence 1, Appli
23	2067	38.8	4002	9	US-09-758-828-3	Sequence 3, Appli
24	2039.5	38.3	3512	10	US-09-749-340-7	Sequence 7, Appli
25	1951	36.6	3861	9	US-09-938-842A-263	Sequence 263, App
26	1868	35.0	2021	10	US-09-873-409-15	Sequence 15, Appli
27	1806.5	33.9	4653	9	US-10-101-388-2	Sequence 2, Appli
28	1698.5	31.9	1940	10	US-09-873-409-16	Sequence 16, Appli
29	1617	30.3	2698	10	US-09-749-340-5	Sequence 5, Appli
30	1538.5	28.9	2681	10	US-09-749-340-8	Sequence 8, Appli
31	1460.5	27.4	3792	9	US-09-882-694-10	Sequence 10, Appli
32	1380	25.9	8036	9	US-10-101-388-1	Sequence 1, Appli
33	1326	24.9	3999	9	US-09-882-694-9	Sequence 9, Appli
34	1101.5	20.7	1175	10	US-09-873-409-11	Sequence 11, Appli
35	1032.5	19.4	7296	10	US-09-070-927A-59	Sequence 59, Appli
36	1002	18.8	9047	10	US-09-070-927A-12	Sequence 12, Appli
37	984.5	17.7	3741	9	US-09-738-626-1251	Sequence 1251, Ap
38	943.5	17.7	12438	10	US-09-070-927A-173	Sequence 173, App
39	902.5	16.9	3601	10	US-09-070-927A-255	Sequence 255, App
40	901.5	16.9	6415	10	US-09-070-927A-275	Sequence 275, App
41	868.5	16.3	640681	10	US-09-790-988-1	Sequence 1, Appli
42	836.5	15.7	1810	10	US-09-749-340-4	Sequence 4, Appli
43	758.5	14.2	2298	9	US-10-156-239-18	Sequence 18, Appli
44	758.5	14.2	2298	10	US-09-795-693-18	Sequence 18, Appli
45	758.5	14.2	3408	9	US-10-156-239-15	Sequence 16, Appli

#### ALIGNMENTS

RESULT 1  
US-09-873-409-12  
; Sequence 12, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 3177  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (198)..(198)  
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)  
US-09-873-409-12

Alignment Scores:  
Pred. No.: 0  
Score: 5328.00  
Percent Similarity: 100.00%  
Length: 3177  
Matches: 1058  
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.96% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-4 (1-1058) x US-09-873-409-12 (1-3177)

Qy	1	MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla	20
Db	1	ATGGTCATCTCATTTGACCAAGTAAGTAAGTCCAAAGCTGGGCTGTGGCA	60
Qy	21	GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu	40
Db	61	GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCGAGGAAAGAACTT	120
Qy	41	GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu	60
Db	121	CAAAAGGCTTTCCCTTTAAATATACCAAGATATGCTGGTTTATTTTCCCCAGTGGCTA	180
Qy	61	LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe	80
Db	181	CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAGATGCAAGGATTTT	240
Qy	81	GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn	100
Db	241	GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTACTTCTTTATGAAT	300
Qy	101	GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly	120
Db	301	GGAACTTATGGACTTGCTTTTGTGTATGGAACCTCCTTGATTTCTTAAGAGAACCTGGGA	360
Qy	121	TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle	140
Db	361	TATACCATCGGACTGTCTTGCTGTTTCTTTAGTGTATCCATAGCAGTATTGCATT	420
Qy	141	GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle	160
Db	421	GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTCATATT	480
Qy	161	PheGlnValIleAspLysLeuProSerIleAspAsnPheSerThrAlaGlyTyrLysPro	180
Db	481	TTCCAGGTTATTGTATAGAAACCCAGATATAGATAACTTTTCCAGCTGGATATAAACCT	540
Qy	181	GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro	200
Db	541	GAATCCATAGAGGACGTGTGGATTTAAATATGTTCTTTCAATTATCCATCAAGACCA	600
Qy	201	SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu	220
Db	601	TCTATCAAGATTCTGAAAGGTCTGAAATCTCAGAATTAAAGTCTGGAGAGACAGTCGCCTG	660
Qy	221	ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp	240
Db	661	GTCCGTCTCAATGGCAGTGGGAAGACGATACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT	720
Qy	241	ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis	260
Db	721	CCGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTCGGCAT	780
Qy	261	TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer	280
Db	781	TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCTGTTTGTGTTCCGGACCCACCATCAGT	840
Qy	281	AsnAsnIleLysTyrGlyArgAspValThrAspGluLeuMetGluArgAlaAlaArg	300
Db	841	AACAATATCAAGTATGGACGAGATGATGTGATGATGAAGAGATGGAGAGACGCAAGG	900
Qy	301	GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly	320
Db	901	GAACAAATGCGTATGATTTTATCATGGAGTTTCTATATAAATTTATATCATGTTGAGG	960
Qy	321	GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu	340
Db	961	GAATAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA	1020

Qy	341	ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer	360
Db	1021	GTTTCAAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCGCTCGGATTCAGAAAGC	1080
Qy	361	LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal	380
Db	1081	AAGTCAGCTGTTCAAGCTGCACTGGAGAGGCGAGCAAGGTCGAGTCAAAATCGTGGTA	1140
Qy	381	AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet	400
Db	1141	GCACACCGACTTCTTACTATTTCGAAGTGCAGATTTGATTGTGACCTTAAAGGATGGAAT	1200
Qy	401	LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu	420
Db	1201	CTGGCGAGAAAGAGGACATGCTGCACTTAATGCAAAACAGAGTCTATATTATTCACTT	1260
Qy	421	ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr	440
Db	1261	GTGATGTCA CAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT	1320
Qy	441	GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle	460
Db	1321	GAAGAAGAACCAACTCACTTCTCTGCACTGTGGAAGAGCATCAAGTCAGACTTCATT	1380
Qy	461	AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu	480
Db	1381	GACAGCTGAGGAATCCACCNAATCTAAAGATTAAGTCTTCTGAAAGTCTCTCTATTTA	1440
Qy	481	LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer	500
Db	1441	AAAAATTTAAAGTTAAACAAGCCTGAATGCGCTTTGTGCTTCTGGGAGCATTTGCTTCT	1500
Qy	501	ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet	520
Db	1501	GTTCTAAATGGAACTGTTCATCCAGTATTTCATCATCTTTGCAAAATTAATAACCATG	1560
Qy	521	PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe	540
Db	1561	TTTGGAATATATATTAACCAACCATTAAGCATGATGCAGAAATTTATTCATGATATTC	1620
Qy	541	ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg	560
Db	1621	GTCAATTTGGTGTATTGCTTTGTGTCAGTTATTCATGCAGGATTTATTTACGGCAGA	1680
Qy	561	AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln	580
Db	1681	GCAGGGGAAATTTTAACGATGAGATTAAAGACACTTGGCCTTCAAAAGCCATGTTATATCAG	1740
Qy	581	AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla	600
Db	1741	GATATTGCTGGTTGATGAAAGGAAACAGCACAGGAGCTTGACACAAATATTAGACC	1800
Qy	601	IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn	620
Db	1801	ATGATATATACAAATTTCAAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAT	1860
Qy	621	AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe	640
Db	1861	GCAACTAACATGGGACTTTTCAGTTATCATTTCTTATATATGATGGAGATGACATTC	1920
Qy	641	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	660
Db	1921	CTGATTTCTGATTTGCTCCAGTACTTGGCGTACAGGAATGATTTGAAACCGCAGCAATG	1980
Qy	661	ThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGlu	680
Db	1981	ACTGGATTTCCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAA	2040
Qy	681	AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet	700
Db	2041	GCTTTGGAGAAATATACGTACTATAGTGTCTATTAAACAGGGGAAAGCCTTCGAGCAATG	2100

QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720  
Db 2101 TATGAAGAGATGCTTTCAGATCTACACAGAAATACCTCGAAGAACACAGATTTATTTGGA 2160  
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740  
Db 2161 AGCTGTATGCATTCAGCCATGCTTTATATATTTTCCCTATGCAGCAGGGTTTCGATTT 2220  
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760  
Db 2221 GGAGCCTATTATTTCAAGCTGGACGAATGACCCACAGAGGCGATGTTTCATAGTTTACT 2280  
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780  
Db 2281 GCAATTGCATATGAGCTATGGCCATCGGAAAAACGCTGCTTTGGCTCCTGAATATTC 2340  
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800  
Db 2341 AAGCCAAATCGGGGCTGGCATCTGTTTGCCTTGTGGNAAAGAAACCAATATATAGAC 2400  
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820  
Db 2401 AGCGCAGTCAAGAAGGCAAAAGCCAGACACATGTCAAGGGAATTTAGAGTTTCGAGAA 2460  
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840  
Db 2461 GTCTCTTTCTTCTATCCATGTCGCCACAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 2520  
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860  
Db 2521 ATTGAGCGAAGAACAGATGACGATTTGTGGGAGCAGCGCTGTGGAAAGACACTTCT 2580  
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880  
Db 2581 GTTCAACTCTGCAGAGACTTATGACCCCGTCAAGACAGTGCCTGTTGATGGGTG 2640  
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900  
Db 2641 GATGCAAAAGAAATTAATGTATACAGTGGCTCGCTTCCCAAAATAGCAATCGTTCTCAAGAG 2700  
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920  
Db 2701 CCTGTGCTCTTCAACTCGAGCATTTGTGAAGAACATCCCTATGTGTGACACAGCCGGTGT 2760  
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940  
Db 2761 GTGCCATTAGATGAGATCAAGAGCCGCAATGCACCAATATCCATTTCTTTATTTGAA 2820  
QY 941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960  
Db 2821 GGTCTCCCTGAGAAATACACACACAAGTTGGACTGAAGAGGAGCACAGCTTTCTGGCGGC 2880  
QY 961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980  
Db 2881 CAGAAACAAAGACTAGTATTGTCAAGGGCTCTCTCCAAAACCCAAAATTTTATTGTTG 2940  
QY 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000  
Db 2941 GATGAGCCCACTTCAGCCCTCGAATATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGAT 3000  
QY 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020  
Db 3001 AAAGCCAGGACGGGAAGGACATGCTAGTGTGCTCACTCACAGGCTCTCTGCAATTCAGAAC 3060  
QY 1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGlu 1040  
Db 3061 GCAGATTGATAGTGGTTCGCAAAATGGAAGATTAAGAAACCAAGAACTCATCAAGAG 3120  
QY 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058  
Db 3121 CTCTCGAANAATCGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 2

US-09-873-409-13  
; Sequence 13, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 3702  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (723)..(723)  
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)  
US-09-873-409-13  
Alignment Scores:  
Pred. No.: 0 Length: 3702  
Score: 5328.00 Matches: 1058  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.96% Indels: 0  
DB: 10 Gaps: 0  
US-09-873-409-4 (1-1058) x US-09-873-409-13 (1-3702)  
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20  
Db 526 ATGGTCACTCTCATTGACCATGAGTAAGTAATTAAGTGCCTATTCGAAGCTGGGGCTGTGGCA 585  
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40  
Db 586 GAAGAAGTCTTGTCATCAATCCGACAGTCATAGCCTTTAGGGCCCGAGGAGAAAGAACTT 645  
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60  
Db 646 CAAAGGTCTTTCCCTTTTAAATATAACAAGATATGCTGGTGGTTTATTTTCCCCAGTGGCTA 705  
QY 61 LeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80  
Db 706 CTAAAGTGTGTTCTGTNTTTTGAAGGTATACAGAACTCTCAAGATGCAAGGATTTT 765  
QY 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100  
Db 766 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTCTGTACTTCTTTATGAAT 825  
QY 101 GlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120  
Db 826 GGAACCTATGGAATGCTTTTGGTATGGAACCTCTCTGATTTCTTAATGAGAACCTGGA 885  
QY 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140  
Db 886 TATACCATCGGACTGTCTTGTCTGTTTCTTTAGTGTAAATCCATGACGATTTATGCATT 945  
QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
Db 946 GGAGCAGCAGTCCCTCACITTTGAACCTTCGCAATAGCCCGAGGAGCTGCTTTCATATT 1005  
QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180  
Db 1006 TTCAGGTTATTGATAAGAAACCCAGTATAGATAAATCTTTTCCACAGCTGATATAAACCT 1065  
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200  
Db 1066 GAATTCATAGAAGGAATCTGGGAATTTAAAAATGTTTCTTTTCAATTATTCATCAAGACCA 1125



QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
Db 1126 TCTATCAAGATTCTGAAAGGCTCGAATCTCAGAAATTAAAGTCGGAGAGACAGTCGCCTTG 1185  
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240  
Db 1186 GTCGGTCTCAATGGCAGTGGGAAAGACGAGTACGGTAGCTCCAGCTTCTGCGAGAGGTATATGAT 1245  
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260  
Db 1246 CCGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCAT 1305  
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280  
Db 1306 TATCGAGACCATAATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGT 1365  
QY 281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300  
Db 1366 AACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGACGAAGG 1425  
QY 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320  
Db 1426 GAACAAATGCGTAGATTTTATCATGGAGTTTCTTAATAAATTTAATACATTTGGTAGGG 1485  
QY 321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340  
Db 1486 GAAAAGGAGCTCAATCAGTGGAGGCGAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1545  
QY 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360  
Db 1546 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1605  
QY 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380  
Db 1606 AAGTCAGCTGTTCAGCTGCACTGGAGAGCGAGAAAGTCGGACTACAAATCGTGGA 1665  
QY 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400  
Db 1666 GCACCCGACTTCTACTATTTCGAAGTCAGATTGTGACCCCTAAAGGATGGAATG 1725  
QY 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420  
Db 1726 CTGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACAGAGTCTATATTATTCATT 1785  
QY 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440  
Db 1786 GTGATGTACAGGATATTAAAGGATGATGAACAGATGGAGTCAATGACATATCTACT 1845  
QY 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460  
Db 1846 GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGACATCAAGTCAGACTTCATT 1905  
QY 461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480  
Db 1906 GACAAGGCTGAGGAATCCACCCTAATCAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1965  
QY 481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500  
Db 1966 AAAATTTTAAAGTTTAAACAAGCCCTGAATGGCTTTTGTGGTTCCTGGGACATTCGCCTCT 2025  
QY 501 ValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThrMet 520  
Db 2026 GTTCTAAATGAACTGTTTCATCCAGATTTTCCATCATCTTTGCGAAAATTTAATACCATG 2085  
QY 521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540  
Db 2086 TTTGGAATAATATGATAAAACACATTAAGCATGATGACAGAAATTTATTCATGATATTC 2145  
QY 541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560  
Db 2146 GTCATTTGGGTGTATTGTGCTTGTGCTAGTTATTTTCATGCGAGGATATTTTACGGCAGA 2205  
QY 561 AlaGlyGluIleLeuThrThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580

Db 2206 GCAGGGGAAATTTTAAACGATGAGATTAAAGACACTTGGCCCTTCAAGCCCATGTTATATCAG 2265  
QY 581 AspIleAlaIleArgPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600  
Db 2266 GATATTGCCCTGGTTTGTATGAAAAGGAAAACAGCACAGGAGGCTTGCAACAATATTAGCC 2325  
QY 601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620  
Db 2326 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTCGCGTCTTAAACACAAAAT 2385  
QY 621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 640  
Db 2386 GCAACTAACATGGCACTTTTCAGTTATCATTTCTTATATATGATGGATGGAGATGACATTC 2445  
QY 641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660  
Db 2446 CTGATTCGTAGTATTGCTCCAGTACTTGGCGTGACAGGAATGATTGAAACCGCAGCAATG 2505  
QY 661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680  
Db 2506 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAA 2565  
QY 681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700  
Db 2566 GCTTTTGAGAAATATACGTACTATAGTGTCAATTAAACAAGGGAAAAGCCTTCGAGCAAAATG 2625  
QY 701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720  
Db 2626 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGA 2685  
QY 721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740  
Db 2686 AGCTGTATGCACTTCAGCCATGCTTTATATATATTTTGCCTATGCAGAGGGTTTCGATTT 2745  
QY 741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760  
Db 2746 GGAGCCTATTTAATCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTTCAT 2805  
QY 761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780  
Db 2806 GCAATTCATATGGAGCTATGGCCATCGGAAAACCGTCGTTTTGGCTCTCGAATATTC 2865  
QY 781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAsp 800  
Db 2866 AAGCCAAATCGGGGCTGCGCATCTGTTTGCCTTGTGGAAAAGAAACCAATATAGAC 2925  
QY 801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820  
Db 2926 AGCCGCACTCAAGAAAGGAAAAAACCCAGACACATGTGAAGGAAATTTAGAGTTTCGAGAA 2985  
QY 821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840  
Db 2986 GTCTCTTCTTCTTCATCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGT 3045  
QY 841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860  
Db 3046 ATTAGCGGAGAAAGACATGATGATTTTGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 3105  
QY 861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880  
Db 3106 GTTCAACTCTGCGAGAGACTTTATGACCCCGTCAAGGACAAAGTGTGTTTGTATGGTGTG 3165  
QY 881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900  
Db 3166 GATGCAAAAAGATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCTCAAGAG 3225  
QY 901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920  
Db 3226 CCTGTGCTCTTCACTGCAGCATTTGTGAGAAACATCGCCTATGTTGACACACGCCGTGTG 3285  
QY 921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940

```
Db 3286 GTGCCATTAGATGAGATCAAGAAGCGCAAAATGCAGCAATATCCATTCTTTATTGAA 3345
Qy 941 GlyLeuProGluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
Db 3346 GGTCTCCCTCAGAAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGG 3405
Qy 961 GlnLysGlnArgLeuAlaLeuAArgAlaLeuGlnLysProLysIleLeuLeu 980
Db 3406 CAGAAACAAGACTAGCTATTGCAAGGGCTCTCTCCAAAACCCAAAATTTATTGTTG 3465
Qy 981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
Db 3466 GATGAGGCCACTTCAGCCCTCGAATATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGAT 3525
Qy 1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
Db 3526 AAAGCCAGGACGGGAAGGACATCGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3585
Qy 1021 AlaAspIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
Db 3586 GCAGATTGATAGTGGTTCTGCACAATGGAAGATTAAGGAACAGGAACCTCAAGAG 3645
Qy 1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
Db 3646 CTCCTGAGAAATCGACATATATTTTAAGTTAGTGAATCGACAGTCAGTGCAG 3699

RESULT 3
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.: 0 Length: 3621
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 96.80% Indels: 27
DB: 10 Gaps: 1

US-09-873-409-4 (1-1058) x US-09-873-409-14 (1-3621)
Qy 1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db 526 ATGGTCATCTCATGACAGTAAGAAATTAAGTCCCTATTTCCAAAGCTGGGGCTGTGGCA 595
Qy 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db 586 GAAGAAAGTCTTGTCATCAATCCGAACAGTCATAGCCCTTTAGGGCCCGAGGAAGAAGACTT 645
Qy 41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
Db 646 CAA----- 648
Qy 61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
Db 649 -----AGTATACACAGAAATCTCAAGATGCAAGAGGATTTT 694
Qy 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
```

```
Db 685 GGCAATAAAAGGACTATAGCTTCAAAAAGTGTCTCTGGGTGCTGTGTACTTCTTTATGAAT 744
Qy 101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db 745 GGAACCTATAGGACTTGTCTTTTGGTATGGAACCTCTCTTGTATCTTAATGGAGAACCTGGA 804
Qy 121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db 805 TATACCATCGGACTGTCTCTGTCTTTCTTTAGTGTAAATCCATAGCAGTATTATTGCATT 864
Qy 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db 865 GGAGCAGCAGTCCCTCCTTCAACCTTTCGAATAGCCGAGGAGCTGCTTTCATATT 924
Qy 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db 925 TTCAGGTTATTGATTAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 984
Qy 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db 985 GAATCCATAGAAGAACTGTGGAATTTAAAAATGTTTCTTCAATTATCCATCAAGACCA 1044
Qy 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db 1045 TCTATCAAGATTCTGAAGGTCTGTAATCTCAGAAATTTAAGTCTGGAGAGACAGTCGCTTG 1104
Qy 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db 1105 GTCCGCTCAATGGCAGTGGAGAGTAGCTAGTCCAGCTTCTGCAGAGTTTATATGAT 1164
Qy 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db 1165 CCGGATGATGGCTTTTATCATGTGGTGAAGATGACATCAGAGCTTTAAATGTCCGCAT 1224
Qy 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db 1225 TATCGAGACCATATTGGAGTGGTAGTCAAGAGCCCTGTTTGTTCGGGACCCACATCAGT 1284
Qy 281 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db 1285 AACATATCAAGTATGGACAGATGATGTGACTGATCAAGAGATGGAGAGACGACGAGG 1344
Qy 301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db 1345 GAAGCAATATGCGTATGATTTTATCATGGAGTTTCTCTAAATAATTTAATACATTTGGTAGG 1404
Qy 321 GluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db 1405 GAAAAGGAGAGCTCAATGATGAGTGGAGGCGAGAAACAGAGGATCGCAATTGCTCGTCTTA 1464
Qy 341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db 1465 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTTACCTCTGCCCTGGATTTCAGAAAGC 1524
Qy 361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db 1525 AAGTCAGCTGTTCAAGCTGCATCTGGAGAGGCGAGCAAAAGGTCGGACTACATCTGTGTA 1584
Qy 381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db 1585 GCACACCGACTTCTTACTATTTCGAGTGCAGATTTGATTTGATGTCACCCCTAAAGGATGAAATG 1644
Qy 401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db 1645 CTGGCGGAGAAAGGAGCACATCTGAACTTAATGGCAAAACGAGGCTATATTATTACATT 1704
Qy 421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db 1705 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATCGAGTCAATGACATATTCTACT 1764
Qy 441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerPhePheIle 460
```

Db	1765	GAAAGAAAGCAACTCACTTCTCTGCGACTCTGTGAAGAGCATCAAGTCAGACTTCATT	1821
Qy	461	AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu	480
Db	1825	GACAGGCTGAGGAATCCACCCAACTCTAAGAGATAAGTCTTCTGAAGTCTCTCTATT	1884
Qy	481	LysIleLeuLysLeuAsnLysProGluTrpPropheValValLeuGlyThrLeuAlaSer	500
Db	1885	AAATTTTAAAGTTAAACAAGCGCTGAATGGCTTTTGTGGTTCTGGGACATTCGCTCTCT	1944
Qy	501	ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet	520
Db	1945	GTTCCTAAATCGAACTGTTTCATCCAGTATTTTCCATCATCTTTTGCAAAAATTTATTAACCATG	2004
Qy	521	PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe	540
Db	2005	TTTGGAAATTAATGATTAACACCAATTAAGCATGATGACAGAAATTTATTCCATGATATTC	2064
Qy	541	ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg	560
Db	2065	GTCAATTTTGGGTGTTATTTTGTCTTTGTCACTATTTTCATCGAGGATTTATTTTACGGCAGA	2124
Qy	561	AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln	580
Db	2125	GCAGGGAAATTTTAAACGATGAGATTAAGACACTTTGGCTTCTCAAGCCATGTTATATACG	2184
Qy	581	AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla	600
Db	2185	GATATTCGCTTGGTTTGATGAAAGGAAACAGCACAGGAGCTTGACAACTATTTAGCC	2244
Qy	601	IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn	620
Db	2245	ATAGATATAGCACAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAACACAAAT	2304
Qy	621	AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe	640
Db	2305	GCAACTAACTGGGACTTTTCAGTTATCATTTCTTTATATATGATGGATGGGAGATGACATTC	2364
Qy	641	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	660
Db	2365	CTGATTTCTGAGTATTTGCTCCAGTACTTGGCGTGACAGGAATGATTTGAAACCGCAGCAATG	2424
Qy	661	ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu	680
Db	2425	ACTGGATTTGCCAACAAAGATAGCAAGAACTTAAAGCATGCTGGAAAGCATAGCAACTGAA	2484
Qy	681	AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet	700
Db	2485	GCTTTGGAGATAATACGTACTATAGTGTCTATTAACAAGGGGAAAAAGCCCTTCGAGCAATG	2544
Qy	701	TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly	720
Db	2545	TATGAAGAGATGCTTTCAGTCTCAACACAGAAATACCTTCGAAAGAAAGCACAGATTATTGGA	2604
Qy	721	SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe	740
Db	2605	AGCTGTTATGCAATTCACCCATGCTTTATATATTTTGGCTATGCGACGAGGTTTCGATTT	2664
Qy	741	GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr	760
Db	2665	GGAGCCTATTATTTCAAGCTGGACGAATGACCCACAGAGGCGATGTTCATAGTTTTTACT	2724
Qy	761	AlaIleAlaTyrGlyAlaMetAlaIleGlyThrLeuValLeuAlaProGluTyrSer	780
Db	2725	GCAATTCGCATATGGAGGTATGGCCATTCGGAAAAACGCTCGTTTTGGCTTCCTGGAATATTC	2784
Qy	781	LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAsp	800
Db	2785	AAAGCCAAATCGGGGGCTGGCACTGTGTTTGGCTTCTTGGAAAAGAAACCAATATAGAC	2844
Qy	801	SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu	820
Db	2845	AGCCGCACTCAAGAGGGAAGAAAGCCAGACACATGTGAAGGAAATTTAGAGTTTCGAA	2904

Qy	821	ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer	840
Db	2905	GTCTCTTTCTTATCCATGTGGCCAGATGTTTTCATCTCCGTGGGTATATCCCTCAGT	2964
Qy	841	IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer	860
Db	2965	ATTGAGCCAGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACTTCT	3024
Qy	861	ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal	880
Db	3025	GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGGACAAAGTCTCTGTTGATGGTGTG	3084
Qy	881	AspAlaLysGlnLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu	900
Db	3085	GATCAAAAGAAATGGAATGTACAGTGGCTCCGTTCCAAATAGCAATGTTCTCTCAAGG	3144
Qy	901	ProValLeuPheAsnCysserIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal	920
Db	3145	CCTGTGCTTCAACTGCAGCATTTCTGAGAACATCGCCTATGTTGCACACAGCCGTGTG	3204
Qy	921	ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu	940
Db	3205	GTGCATTTAGATCAGATCAAAAGAGCCGAAATGCAGCAAAATATCCATCTTTTATTGAA	3264
Qy	941	GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyValaGlnLeuSerGlyGly	960
Db	3265	GGTCTCCTCGAANAATACACACACAACTTGGACTGAAAGGACACAGCTTCTGGCGGC	3324
Qy	961	GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu	980
Db	3325	CAGAAACAAAGACTAGCTATTGCAAGGGCTCTCTTCCAAAAACCCAAAAATTTATTGTTG	3384
Qy	981	AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp	1000
Db	3385	GATGAGGCCACTTCAGCCCTCGAATAATGACGTGAGAAGGTGTTTCAGCATGCCCTTGAT	3444
Qy	1001	LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn	1020
Db	3445	AAAGCCAGGACGGGAGGACATGCTTAGTGGTCACTCA CAGGCTCTCTGCAATTCAGAAC	3504
Qy	1021	AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu	1040
Db	3505	GCAGATTGTAGTGGTTCGTGCACATGGAAAGATAAAGAACAGAGGAACCTCATCAAGAG	3564
Qy	1041	LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln	1058
Db	3565	CTCTCGAATTCAGACATATATTTTAACTTAGTGAATGCAATGCACAGTCACTGACAG	3618

## RESULT 4

```

US-09-873-409-10
; Patent No. 10, Application US/09873409
; Sequence 10, Application US/09873409
; General Information:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding
; TITLE OF INVENTION: Homologue on C
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873-409-10
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 10
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-10

```

Alignment Scores:	
Pred. No.:	0
Score:	4561.50
Percent Similarity:	97.96%
Length:	2856
Matches:	912
Conservative:	0

Best Local Similarity: 97.96%			Matches: 1		
Query Match: 85.58%			Indels: 18		
DB: 10			Gaps: 1		
US-09-873-409-4 (1-1058) x US-09-873-409-10 (1-2856)					
QY	129	ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGI	148		
DB	114	GTATTCTTTAGTGAATCCATAGCAGTATTTCATTTGGAGCAGCAGTCCCTCATTATT--	171		
QY	148	uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr	168		
DB	172	-----GATAAGAAACC	182		
QY	168	oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGI	188		
DB	183	CAGTATAGATAAATTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGTGGA	242		
QY	188	uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe	208		
DB	243	ATTATAAAATGTTTCTTCAATATATCCATCAAGACCATCTATCAAGATTCTGAAGGTCT	302		
QY	208	uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy	228		
DB	303	GAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGGCAGTGGGAA	362		
QY	228	sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa	248		
DB	363	GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT	422		
QY	248	lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa	268		
DB	423	GGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCATTCAGACCATATTCGAGTGGT	482		
QY	268	lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs	288		
DB	483	TAGTCAAGAGCCCTGTTTGTTCGGGACCCATCAGTAACAATATCAAGTATGGACGAGA	542		
QY	288	pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheII	308		
DB	543	TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAGCAAGGGAAGCAATTCGATGATTTAT	602		
QY	308	eMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGI	328		
DB	603	CATGGAGTTTCTTAATAAATTAATACATGTTAGGGGAAAAGGAGCTCAAAATGAGTGG	662		
QY	328	yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuII	348		
DB	663	AGGGCAGAAACAGAGGATCGCAATTCGCTGCTGCTTAGTTTGGAAACCCCAAGATTCTGAT	722		
QY	348	eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe	368		
DB	723	TTTAGATGAGGCTACGCTGCTCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCAT	782		
QY	368	uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr	388		
DB	783	GGAGAAAGCGAGCAAAAGGTCCGACTACAATCGTGTAGCACCCAGCTTCTTACTATTTCG	842		
QY	388	gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl	408		
DB	843	AAGTGCAGATTGATTTGTGACCTTAAAGGATGGAATGCTGCGCGAGAAAGGAGCAGCATGC	902		
QY	408	aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy	428		
DB	903	TGAACCTAATGGCAAAACGAGGTCTATATTATTCATTTGTGATGTCACAGGATATTAAAAA	962		
QY	428	sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr	448		
DB	963	AGCTGATGAACAGATGGAGTCAATGACATATTCATCTACTGAAAGAAAGACCAACTCACTTCC	1022		
QY	448	oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGI	468		
DB	1023	TCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTTCATTGACAAGGCTGAGGAATCCACCCA	1082		
QY	468	nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr	488		
DB	1083	ATCTAAAGAGATAAGTCTTCTCCTGAAGTCTCTCTATTAAAAAATTTAAAGTTAAACAAGCC	1142		
QY	488	oGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr	508		
DB	1143	TGAATGGCTTTTGTGGTCTTGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTTATCC	1202		
QY	508	oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh	528		
DB	1203	AGTATTTCATCATCTTTGCAAAAATTAATACCATGTTTGGAAATATATGATAAACCCAC	1262		
QY	528	rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh	548		
DB	1263	ATTAAAGCATGATGCAGAAATTTATTCATGATATTCGTCATTTTGGGTGTTATTGCTT	1322		
QY	548	eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr	568		
DB	1323	TGTCAGTTATTTTCATGCGGGATTATTTTACGCGACAGAGGGGAAATTTTAAACGATGAG	1382		
QY	568	gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLy	588		
DB	1383	ATTAAGACACTTGGCCCTTCAAGCCATGTTATATCAGGATATTCCTGTTTCATGNAAA	1442		
QY	588	sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGI	608		
DB	1443	GGAAAAACAGCACAGGAGGCTTGACAAATATTTAGCCATAGATATAGCAAAATTTCAAGG	1502		
QY	608	yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa	628		
DB	1503	AGCAACAGGTTCCAGGATTTGGCGTCTTAAACAAAAATGCAACTTAACATGGGACTTTTCAGT	1562		
QY	628	lIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVa	648		
DB	1563	TATCATTTCTTTATATATGATGGAGGAGATGACATTCCTCGATTCTCGATATTGCTCCAGT	1622		
QY	648	lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy	668		
DB	1623	ACTTGGCGTGACAGGAATGATTGAAACCCGACGCAATGACTGGATTTCGCCAACAAAGATAA	1682		
QY	668	sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrII	688		
DB	1683	GCAAGAACTTAAGCATGCTGGAAAGATAGCACTGAAGCTTTTGGAGAAATATACGTACTAT	1742		
QY	688	eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGI	708		
DB	1743	AGTGTCAITTAACAGGGAAGAACCTTCGAGCAAAATGTTATGAAGAGATGCTTCAGACTCA	1802		
QY	708	nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl	728		
DB	1803	ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC	1862		
QY	728	aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGI	748		
DB	1863	CTTTATATATTTTGGCTATGTCAGCAGGGTTTCGATTGAGGCTTATTTAATTCAGGTGG	1922		
QY	748	yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl	768		
DB	1923	ACGAATGACCCACAGAGGATGTTTCATAGTTTTCGCAATTCGCATATGAGACTATGGC	1982		
QY	768	alleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi	788		
DB	1983	CATCGAANAACCGCTGTTTGGCTCTGATATTTCCAAAGCCCAATTCGGGGCTCGGCA	2042		
QY	788	sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy	808		
DB	2043	TCGTGTTGCTTCTTGAAAAAGAAACCAATATAGACAGCCGCGAGTCAAGAGGGGAAAAA	2102		
QY	808	sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr	828		
DB	2103	GCCAGACACATGTGAAGGGAAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG	2162		

```
QY 828 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 848
DB 2163 CCCAGATGTTTTCATCTCCGTCGCTTATCCCTCAGTATTGAGCGAGAAAGACAGTAGC 2222
QY 848 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 868
DB 2223 ATTTGTGGGAGCAGCGCGCTGTGGAAAGCAGCTTCTGTTCAACTTCCTGAGAGACTTTA 2282
QY 868 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValG1 888
DB 2283 TGACCCCGTCGACAGACAGTCTGTTGATGTGTGATGTCGAAAGAAATGATGTACA 2342
QY 888 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 908
DB 2343 GTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCGCTGTGCTCTTCAACTGCAGCAT 2402
QY 908 eAlaGluAsnIleAlaTyArgAspAsnSerArgValValProLeuAspGluIleLysG1 928
DB 2403 TGCTGAGAACATCGCTATGTGACAAACAGCGCTGTGTGTCGATAGATGAGATCAAGA 2462
QY 928 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyArgnTh 948
DB 2463 AGCCGCAATGCAGCAATATCCCAATCTTTTATTGAAGGCTCTCCCTGAGAAATACAACAC 2522
QY 948 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 968
DB 2523 ACAAGTTGGACTGAAGGAGCACAGCTTCTGGCGGCAGAAACAAAGACTAGCTATTGC 2582
QY 968 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 988
DB 2583 AAGGCTCTTCTCCAAAACCCAAAATTTTATTGTGATGAGGCCACTTCAGCGCCTCGA 2642
QY 988 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1008
DB 2643 TAATGACAGTGAGAGGTTGGTTTCAGCATGCCCCCTTGATAAAGCCAGGAGGAGACATG 2702
QY 1008 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1028
DB 2703 CCTAGTGTCTACTCACAGGCTCTCTGCAATTCAGAACGCGAGATTGATGTTCTGCA 2762
QY 1028 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1048
DB 2763 CAATGGAAGATAAAGGAACAAGGAACCTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822
QY 1048 rPheLysLeuValAsnAlaGlnSerValGln 1058
DB 2823 TTTTAAGTTAGTGAATGCACAGTCAGTCAG 2853
```

## RESULT 5

```
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-9

Alignment Scores:
Pred. No.: 3,38e-314 Length: 2066
Score: 3456.00 Matches: 687
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 64.84% Indels: 0
DB: 10 Gaps: 0
US-09-873-409-4 (1-1058) x US-09-873-409-9 (1-2066)
QY 372 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 391
DB 3 AGCAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTTCTACTATTCCAGAGTCAGAT 62
QY 392 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 411
DB 63 TTGATTGTGACCTTAAGGATGGAATGCTGGCGAGAAAGGAGCACATGCTGAACATAATG 122
QY 412 AlaLysArgGlyLeuTyTrpSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 431
DB 123 GCAAAACGAGGCTATATTATTACCTTGATGTACAGGATATTAAAAAGCTGATGAA 182
QY 432 GlnMetGluSerMetThrTyTrpSerThrGluArgLysThrAsnSerLeuProLeuHisSer 451
DB 183 CAGATGGAGTCAATGACATATTCTCTGAAAGAAAGACCAACTCCTTCTCTGCACTCT 242
QY 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 471
DB 243 GTGAAGAGCATCAAGTCAGACTTCATTGACAGGCTGAGGAATCCACCCCAATCTAAGAG 302
QY 472 IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrpPro 491
DB 303 ATAAGTCTTCTCAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGCGCT 362
QY 492 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 511
DB 363 TTTGTGGTCTCTGGGACATTTGGCTTCTGTTCTAAATGGAACCTGTTCATCCAGTATTTTC 422
QY 512 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 531
DB 423 ATCATCTTTGCCAAAATTTATTAACATGTTTGGAAATTAATGATAAACCACATTAAGCAT 482
QY 532 AspAlaGluIleTyTrpSerMetIlePheValIleLeuGlyValIleCysPheValSerTy 551
DB 483 GATGCAAAATTTATTCCATGATATTCTGTTTGGGTGTTTATTGCTTTGTCAGTTAT 542
QY 552 PheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 571
DB 543 TTCATGAGGGATTTATTTTACGGCAGAGCGGGAATTTTAAAGATGAGATTAAGACAC 602
QY 572 LeuAlaPheLysAlaMetLeuTyTrpGlnAspIleAlaTrpPheAspGluLysGluAsnSer 591
DB 603 TTGGCCTTCAAGCCATGTTATATCAGGATATTCCTGCTGGTTTGGTGAAGAAAGAACAGC 662
QY 592 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 611
DB 663 ACAGGAGGCTTGCAACAATATTAGCCATAGATATAGCAAAATTCAGGAGCAACAGGT 722
QY 612 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 631
DB 723 TCCAGGATTTGGCGTCTTAAACACAAAATGCAACTAACATGGGACTTTTTCAGTTATCATTTCC 782
QY 632 PheIleTyTrpGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 651
DB 783 TTTATATATGATGGAGATGACATTCCTGATTCCTGAGTATTGCTCCAGTACTTGGCGTG 842
QY 652 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 671
DB 843 ACAGGAATGATTGAACCCGAGCAATGACTGGAATTCCTCAACAAGAAATAGCAAGAACTT 902
QY 672 LysHisAlaGlyLysIleAlaThrGluAlaLeuAsnIleArgThrIleValSerLeu 691
DB 903 AAGCATGCTGGAAGATAGCAACTGAAAGCTTTGAGAAATATACGTACTATATAGTGTCAATTA 962
QY 692 ThrArgGluLysAlaPheGluGlnMetTyTrpGluGluMetLeuGlnThrGlnHisArgAsn 711
DB 963 ACAAGGGGAAAAGGCTTCGAGCAAAATGTTATGAGAGATGCTTTCAGACTCAACACAGAAAT 1022
```





Qy	184	GlulGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	203
Db	1237	AAGGGAAATTTGGAAATTCAAAATGTTCACTTCAGTTACCTCTCTCGAAAGAAGTTAAG	1296
Qy	204	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	223
Db	1297	ATCTTAAAGGCTCTCAACCTGAAGGTTCAAGTGGGCGACAGTGGCGCTGGTTGGGAAC	1356
Qy	224	AsnGlySerGlyLysSerThrValValGlnLeuLeuLeuGlnArgLeuTyrAspProAspAsp	243
Db	1357	AGTGGCTGGCGGAAGACACAGCCGTGCAGCTGATGCAGAGGCTCTATGACCCACAGAT	1416
Qy	244	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	263
Db	1417	GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA	1476
Qy	264	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	283
Db	1477	ATTACTGGTGTGTAGTCAGGAGCCGTGTGTGTGTGGCCACCGATAGCTGTGAACAACATT	1536
Qy	284	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	303
Db	1537	CGCTATGGCCGCGAAATGTCACCATGGATGAGATGAGAAGGCTGTTAGAAGACCCAT	1596
Qy	304	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323
Db	1597	GCCTATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTGTTGGAGAGAGGG	1656
Qy	324	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	343
Db	1657	GCCCGGTGAGTGGTGACAGAACAGAAATCGCCATTTGCTGGGCGCCTGGTTTCGCAAC	1716
Qy	344	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
Db	1717	CCCAAGATTCTTCTGCTGGATGAGGCAAGCTCAGCTCTGGACACTGAAAGTGAAGCAGTG	1776
Qy	364	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	383
Db	1777	GTTCAGGTGSCCTGGATAAGGCCAGAAAGCCGGAGCTACCATTTGTGATAGCTCATCGT	1836
Qy	384	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
Db	1837	TTGTCTACAGTTCTGTAATCCGATGTCATTTGCTGGTGTGTTGATGATGAGTCAITTTGGAG	1896
Qy	404	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	423
Db	1897	AAAGGAATCATGATGAACCTCATGAAGAGAGAGGCGATTTACTTCAAACCTGTGCACATG	1956
Qy	424	GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys	443
Db	1957	CAG--ACAAGAGAAATGAATGAGTTAGAAATGCCACTGGTGAATCCAAAGTGAA	2013
Qy	444	ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle-----	460
Db	2014	AGTGATGCCTTGGAAATG--TCTCAAAGAGATTGAGGTCGAGTTAATAAAGAAAGA	2070
Qy	461	-----AspLysAlaGluGluSerThrGln	468
Db	2071	TCAACTCGCAGGAGTATACATGCACCAAGGCCCAAGACAGAAAGCTTGGTACAAAGAG	2130
Qy	469	SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro	488
Db	2131	GACTTGAATGAGAATGTTCCCTTCAGTTCCTTCTGGAGGATTCGAAGCTGAACCTCAACT	2190
Qy	489	GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro	508
Db	2191	GAATGGGCTTATTTTGTGGTGTGTATATTTGTGTCTATTATAAAGCGAGGCGCTCCACCA	2250
Qy	509	ValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAspLysThr	527
Db	2251	GCATTTTCAATAATATTTTCAAGGATATAGGATCTTTTACCCGAGATGAGATCTCTGAA	2310
Qy	528	ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleLys	547

Db	2311	ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTCTAGTCCCTGGAAATTTATTCTT	2370
Qy	548	PhcValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuThrMet	567
Db	2371	TTTATTACATTTTTCTCCAGGCGTCCATTTGGCAAGCTGGGAGATCCTCACTAAG	2430
Qy	568	ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrrPheAspGlu	587
Db	2431	CGGCTTCGATACATGGTTTCAGATCCATGCTGACACAGGATGCAGCTGGTTGATGAC	2490
Qy	588	LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln	607
Db	2491	CCTAAAAACACCACTCGGACCATTTGCAACACAGGCTTGCCAATGATGGGCTCAAGTTAAA	2550
Qy	608	GlyAlaThrClySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer	627
Db	2551	GGGCGCTATAGTTCCAGGCTCTGCTGATACCCAGAAATATAGCAAAATCTTGGGACAGGC	2610
Qy	628	ValIleIleSerPheIleTyrGlyTrrPleuMetThrPheLeuIleLeuSerIleAlaPro	647
Db	2611	ATTATTATATCTTATCTATGTTGGCAATTAACACATTTTACTCTTAGCAATTTGTACC	2670
Qy	648	ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp	667
Db	2671	ATCATTTGCAATAGACGAGTTCTGAAATGAAATGTTGTCTGACACAGCACTGAAAGAT	2730
Qy	668	LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr	687
Db	2731	AGAAGAAGCTAGAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAANAACCTCCGAAC	2790
Qy	688	IleValSerLeuThrArgLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr	707
Db	2791	GTGTTTCTTTGACTCGGAGCAGAAGTTTGAATACATGATGACACAGAGTTTCCAAGTA	2850
Qy	708	GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis	727
Db	2851	CCATACAGAAACTCTTTGAGGAAGACACACATCTTCGGGGTCTCATTTTCTATCACCAC	2910
Qy	728	AlaPheIleTyrPheAlaTrrAlaAlaGlyPheArgPheGlyAlaTrrLeuIleGlnAla	747
Db	2911	GCAATGATGATTTTTTCCATGCTGGCTGTTTCCGGTTTGGTGCTACTTGGTGCGAAAT	2970
Qy	748	GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTrrGlyAlaMet	767
Db	2971	GAGTTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGCTTTGGTGCCATG	3030
Qy	768	AlaIleGlyLysThrLeuValLeuAlaProGluTrrSerLysAlaLysSerGlyAlaAla	787
Db	3031	GCAGTGGGCGAGGTCAGTTTCATTTGCTGCTACTATGCCAAGCCAAAGTATCAGCAGCC	3090
Qy	788	HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys	807
Db	3091	CAGCTCATCATGATCATTAAGAAAAGCCCTCTGATGTACAGCTACAGCCCTCACGGCCTC	3150
Qy	808	LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCys	827
Db	3151	AAGCCAAATACGTTGGAAGAAATGTGACATTTAATGAGGTGCTGTTCAACTATCCCACT	3210
Qy	828	ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal	847
Db	3211	CGACCAACATCCCGCTGCTCCAGGCGCTGAGCCCTCAGGTCGAGGTGAAGAAGGGCCAGACG	3270
Qy	848	AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu	867
Db	3271	GCCCTGTAGGTAGCATGGCTGTGGGAAGACACAGTTGTTTACGCTCCTAGCGGCTTC	3330
Qy	868	TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal	887
Db	3331	TATGACCCCTTGCTGGTTTCAGTGTCTAATTGATGCCAAGAGATAAAGCACCTGAATGTC	3390
Qy	888	GlnTrrPleuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer	907



Db 3391 CAGTGGCTCGAGCACACCTGGGCATCGTCTCTCAGAGGCCATCTGTTGACTGCAGC 3450  
Qy 908 ILeAlaGluAsnIleLaTyRgLyAspAsnSerArgValValProLeuAspGluIleLys 927  
Db 3451 ATTGCCGAGAACTTGCCTATGTGAGACAAACAGCCGGTCTATCAGATGAAGAGATTATG 3510  
Qy 928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluClyLeuProGluIlyTyRAsn 947  
Db 3511 CAGGCGCCAGAGGCGCAACATACACCTTTCATCGAGACATCTCCCTGAGAAATACAAC 3570  
Qy 948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967  
Db 3571 ACCAGATGAGACAAAGAACCCAGCTCTCTGTTGGCCAGAAACAGCGCATGGCCATA 3630  
Qy 968 AlaArgAlaLeuLeuGlnTyRAspGlyLeuLeuLeuAspGluAlaThrSerAlaLeu 987  
Db 3631 GCTCGCGCTCTGTAGACAGCTCATATTTGCTTTTGGATGAAGCTACATCAGCTCTG 3690  
Qy 988 AspAsnAspSerGluLysValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007  
Db 3691 GATACAGAAAGTGAAAGGTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGCCACC 3750  
Qy 1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027  
Db 3751 TGCAATGTGATGCCCCACCCTGTCACCATCCAGAAATGCAGATTTATATAGTGGTGT 3810  
Qy 1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAspIle 1047  
Db 3811 CAGAATGGCAAGTCAAGAGATGGCACACATCACAGCTGCTGGCTCAGAAAGGCATC 3870  
Qy 1048 TyrPheLysLeuValAsnAlaGlnSer 1056  
Db 3871 TATTTTCCATGATCAGTGTCCAGGCT 3897

## RESULT 7

US-09-769-097-1  
; Sequence 1, Application US/09769097  
; Patent No. US20020055128A1  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Eilens  
; APPLICANT: John Anthony Feild  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; FILE REFERENCE: GP-50009-C2  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR FILING DATE: 1998-12-09  
; PRIOR FILING DATE: 1998-12-09  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR FILING DATE: 1999-09-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 4369  
; TYPE: DNA  
; ORGANISM: RATTUS RATTUS  
US-09-769-097-1

## Alignment Scores:

Pred. No.:	5,7e-281	Length:	4369
Score:	3106.50	Matches:	600
Percent Similarity:	75.54%	Conservative:	206
Best Local Similarity:	56.23%	Mismatches:	218
Query Match:	58.28%	Indels:	43
DB:	10	Gaps:	7

US-09-873-409-4 (1-1058) x US-09-769-097-1 (1-4369)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyRgLySerLysAlaGlyAlaValAlaGluVal 23  
Db 1039 TCATTTTCATGAAGAAGACTCCAGGCTTATGCAAAAGCTGGAGCACTTGTCTGAAGAGTC 1098  
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgSer 43  
Db 1099 TTACAGCCATCAGAACTGTGATTGCTTTGGAGGACAAAAGNAGAACTTGAA----- 1152  
Qy 44 PheLeuLeuAsnIleThrArgTyRAlaTyRPhPheTyRPhPheProGlnTrpLeuLeuSerCys 63  
Db 1152 ----- 1152  
Qy 64 ValLeu\*\*\*PheValArgTyRThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 1153 -----AGGTACAATAAACAATTTGGAAGAGCTAAAGAGGTGGGATAAAG 1197  
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyRPhPheMetAsnGlyTyR 103  
Db 1198 AAGCTATTCAGGCCAACATTTCCATGGGTGAGCTTTTCTGCTTATCTATGTCATCATAT 1257  
Qy 104 GlyLeuAlaPheTyRgLyThrSerLeuIleLeuAsnGlyGluProGlyTyRThrIle 123  
Db 1258 GCTCTGGCATTCGTGATGGACTTCCTTGGTCATCTCAAAAGAA-----TACACTATT 1311  
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyRcysIleGlyAlaAla 143  
Db 1312 GCACAAGTCTCAGCTCTCTTTTCTGTATTAATTTGGAGCATTCAGTGTGGCAGGCA 1371  
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 1372 TCTCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431  
Qy 164 IleAspLysGlyProSerIleAspAsnPheSerThrAlaGlyTyRcysProGluSerIle 183  
Db 1432 ATTGATAATTAAGCCAGTATAGACAGCTTCTCAAGAGTGGGCACAAACCCGACACATA 1491  
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyRProSerArgProSerIleLys 203  
Db 1492 CRAAGAAATTTGGAATTCAAAATATTCACTTCAGTTACCCGTCGCAAAAGACGTTCAG 1551  
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1552 ATCTGAAGGGCTCAACCTGAAGGTGAAGCGGCAGACGGTAGCCCTGGTTGGCAAC 1611  
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyRAspProAspAsp 243  
Db 1612 AGTGGCTGTGGAAAAGACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1671  
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyRArgAsp 263  
Db 1672 GCGCAGGTCAGTATCCAGCGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGAA 1731  
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1732 ATCATTTGGGGTGTGAGTCAGGAACCCGCTGCTTTGCCACCACCAATTCGCGAAAACATT 1791  
Qy 284 LysTyRcysArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1792 CGCTATGGCCGAGAAAACGTCACCATGGATGATAGAGAAAGCTGTCAAGGAGGCAAT 1851  
Qy 304 AlaTyRAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323  
Db 1852 GCCTATGATTTTCATCATGAACCTGCCCCCAAAATTTGACACCTGTTGTTGAGAGGGG 1911  
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1912 GCGCAGCTGAGTGGGGGACAGAAACAGAGATCGCCATTTGCCCGGCCCTGTGGTCCGCAAC 1971  
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysAla 363  
Db 1972 CCCAAGATCTTTTGTGATGAGGCCAGCTCAGCTTGGACACAGAAAGCGGCGGTG 2031  
Qy 364 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383

Db 2032 GTTCAGCCGCTCTGGATAAGCTAGAGAGCGCGACCACTTGTGATGATCACCAGC 2091  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 2092 TTGTCTACAGTTCGCAATGCTGACGTCATTCGCTGTTTGGATGGTGTGTCATG 2151  
Qy 404 LysGlyAlaHisAlaGluMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 2152 CAAGGAATCATGATGATGCTCATGAGAGAAAGAAATTTACTTCAAACTGTGTCAT 2211  
Qy 424 Gln-----AspIleLysAlaAspGluGlnMetGluSer----- 435  
Db 2212 CAGACAGCAGGAAATGAATAGGAATCAAGCTTGTGAATCTAAAGACGGAATT 2271  
Qy 436 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451  
Db 2272 GATAATGTGGACATGCTCTCAAAAGATTCGGGATCCGATCTAATAAGAAAGATCAACT 2331  
Qy 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470  
Db 2332 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGAACCTTAGCACCAAGAGGCTCTG 2391  
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 490  
Db 2392 GATGACGACGTACCTCCAGCTTCCTTTGGCGGATCCTGAAGTTGAATCAACTGAATGG 2451  
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2452 CCTTATTTTGTGGTGTGTTATTTGTGCCATAATAAATGGAGGCTTGCACCAACGATTC 2511  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529  
Db 2512 TCCATAATATTTCAAAGGTTGTAGGGGTTTTTACAAAAATGACACCCCTGGAATCCAG 2571  
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 2572 CGGCAGAACACCAACTGTTTCTTATTGTTCTGTATCTCTGGATCATCTCTTTCAAT 2631  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 2632 ACGTTTTTCTTCAGGGCTTCACATTTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACT 2691  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589  
Db 2692 CGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAGCTGTTGATGACCTTAA 2751  
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2752 AACACCACAGGAGCGCTGACCCACGAGCTTGCCAATGACGCTGCTCAAGTCAAAAGGGCT 2811  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2812 ACAGGCTCTAGGCTTGTGTTATTACCAGAACATAGCAATCTTGGGACAGGCTATC 2871  
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2872 ATATCCCTGATCTAGGCTGCAATGTGACTTTTACTCTTAGCAATTTGTTCCCATCAT 2931  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2932 GCTATAGCAGGAGTGGTTGAATGAAATGTTGCTGGCAAGCGCTGAAAGATAAGAAG 2991  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2992 GAACTAGAAAGTTCGGGAAGATCGCTACAGAAGCAATGAAATCTTTCGCACTGTCGC 3051  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 709  
Db 3052 TCTTTGATCGGGACGAAAGTTTGAATGATGATGCCAGAGCTTCAGATACCATTAC 3111  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729

Db 3112 AGAAATGCTTTGAGAAAGCGCAGCTCTTTGGGATCAGCTTTCTCTTCCACCCAGGCCATG 3171  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValaTyrLeuIleGlnAlaGlyArg 749  
Db 3172 ATGATTTCTCCCTATCTGCTGTTTCCGTTTGTGCTTACTTGTGGCACAGAACTC 3231  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 3232 ATGACATTTGAAATGTTCTGTTAGTATTCTCAGCTATTGTTCTTGTGTCATGGCAGTG 3291  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 789  
Db 3292 GGCAGCTCAGTTTCATTCGCTCTGACTAGCGAAAGCCAAAGCTCTCGGCATCCCCATC 3351  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3352 ATCAGGATCATTTGAGAAATCCCTGAGATTGACAGTACAGCAGCGAGGGCTTGAAGCCT 3411  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 3412 AATATGTTGGAAGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCACCCGACCC 3471  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3472 AACATCCCATGCTTCAGGGGCTGAGCTTAGAGGTGAAGAAAGCGCAGCTGGGCCCTC 3531  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 3532 GTGGGACAGCTGGCTGCGGAAAGATGACAGTGTCCAGTCTTGTGAGCGCTTCTATGAC 3591  
Qy 870 ProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889  
Db 3592 CCATGCGCGGACAGCTGTTCTTAGATGGCAAGAAATAAAGCAACTCAATGTCAGTGG 3651  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3652 CTCGCGGCCACCTGTGGCATTTGTGCCAGAGGCCATCTCTGTTGATCGACATCGCC 3711  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3712 GAGAACATTCCTACGGAGACACAGCCGTGCTGTCTCATAGGAGATCGTGAAGGCA 3771  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3772 GCCAAGAGGCGCAACATCCACCACTCATCGACTCGCTGAGAAATAACAACACCAGA 3831  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3832 GTGGAGACAAAGGACTCAGCTGTGCGGCGGCGAGCAGCGCATCGCCATCGCGCGC 3891  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 3892 GCCCTGCTCAGACAGCTCATCTTACTTCTGATGAGCGACATCAGCTCTGATACG 3951  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 3952 GAGAGTCAAAAGGTCCTCCAGGAAGCGCTGACAAAGCCAGGAAGCCCGCACCTGCATT 4011  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 4012 GTGATCGCAGCGCTGTCCACCATCCAGAACGACAGCTTGATCGGTGATTCAGAAC 4071  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 4072 GCCAGGTCAAGGAGCAGCGGCCACCCACAGACGCTGTGCGGTGATTCAGAAC 4131  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 4132 TCGATGCTCAGTGTGCAAGCT 4152  
RESULT 8  
US-09-769-097-3  
; Sequence 3, Application US/09769097

Patent No. US20020055128A1

GENERAL INFORMATION:

APPLICANT: Kimberly Anne Brun

APPLICANT: Richard James Chenery

APPLICANT: Harma Ellens

APPLICANT: John Anthony Feild

APPLICANT: Lin Yue

TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES

TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF

FILE REFERENCE: GP-50009-C2

CURRENT APPLICATION NUMBER: US/09/769,097

CURRENT FILING DATE: 2001-01-25

PRIOR APPLICATION NUMBER: 09/208,809

PRIOR FILING DATE: 1998-12-09

PRIOR APPLICATION NUMBER: 09/156,800

PRIOR FILING DATE: 1999-09-17

PRIOR APPLICATION NUMBER: US99/20770

PRIOR FILING DATE: 1999-09-10

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 3

LENGTH: 4425

TYPE: DNA

ORGANISM: RATTUS RATTUS

US-09-769-097-3

Alignment Scores:

Pred. No.: 5,81e-281 Length: 4425  
Score: 3106.50 Matches: 600  
Percent Similarity: 75.54% Conservative: 206  
Best Local Similarity: 56.23% Mismatches: 218  
Query Match: 58.28% Indels: 43  
DB: 10 Gaps: 7

US-09-873-409-4 (1-1058) x US-09-769-097-3 (1-4425)

Qy 4 SerLeuThrSerLysGluSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23  
Db 1039 TCATTTACTGATAAGGAACCTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 1098  
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
Db 1099 TTAGCAGCCATCAGAACTGTGATTCGCTTTGGAGGACAAAGAGGAAGTGAATTTGAA 1152  
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaIlePheTyrPheProGlnTrpLeuLeuSerCys 63  
Db 1152 ----- 1152  
Qy 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 1153 -----AGGTACAATAACAATTTGGAAGAAGCTAAAGGCTTGGGATAAAG 1197  
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
Db 1198 AAGCTATACGGCCACATTTTCATGGGTCAGCTTTCTGCTTATCTATGATCATAT 1257  
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrIle 123  
Db 1258 GCTCTGGCATCTGATAGGACTTCTCTGGTCACTCAAAAGNA-----TACACTATT 1311  
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143  
Db 1312 GGCAAGTGCTCACTGCTCTTTTCTGTATTAATTGGAGCATTCAGTGTGGCAGGCA 1371  
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePheGlnVal 163  
Db 1372 TCTCCAAATATGAAGCTTCGCAATTCGTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431  
Qy 164 IleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 1432 ATTGATAATGAAGCCAGTATAGAGCTTCTCAAAAGTGGGCAAAACCCGACACATA 1491  
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203

Db 1492 CAAGGAAATTTGGAATTCAAAATATATTCATCTCAGTTACCCGTCTCGAAAAGACGTTTCAG 1551  
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1552 ATCTTGAAGGGCTCAACCTGAAGGTGAAGAGGGGACAGCGTAGCCCTGGTTGGCAAC 1611  
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
Db 1612 AGTGGCTGTGGGAAAAGCAACATGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1671  
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1672 GCGAGGTCAAGTATCAGCGGACAGGACATCAGACCATCAATGTGAGGTATCTGCGGGA 1731  
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnile 283  
Db 1732 ATCATTTGGGTGTGAGTCAGGAACCCGTGCTTTTCCACCACCAATTCGCGAAAACATT 1791  
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1792 CGCTATGGCCGAGAAACGTCACCATGATGATAGAGAAAGCTGTCAAGGAAGCCAAT 1851  
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323  
Db 1852 GCCTATGATTTTCATCATGAACCTGCCCAAAATTTGACACCTTGTGTGAGAGGG 1911  
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1912 GCGAGCTCAGTGGGGGACAGAAACAGAGAGATCCCATTTGCCGGGCCCTGGTCCGCAAC 1971  
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1972 CCCAAGATCCTTTTGTGATGAGGCCACGTCAGCCTTGACACAGAAAGCGGCGTG 2031  
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 383  
Db 2032 GTTCAGGCCGCTCTGATGAAGCTAGAGAGGCGGACCAACCATTTGTGATGATCTACCCG 2091  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 2092 TTGCTACAGTTCGCAATGCTGATGTCATTTGCTGGTTTGTGCTGGTGTCTATGGAG 2151  
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 2152 CAAGGAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2211  
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 435  
Db 2212 CAGACAGCAGGAATGAAATTTAGGAATTTAGGAATTTAGGAATTTAGGAATTTAGGAATTT 2271  
Qy 436 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeuHisSer 451  
Db 2272 GATAATGTGGACATGCTTCAAAAAGATTCCGGATCCAGTCTAATAAGAGAAATCAACT 2331  
Qy 452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470  
Db 2332 CGCAAAAGCATCGTGGGCCACATGATCAAGACGGGAACTTAGCACCAAGAGGCTCTG 2391  
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTrp 490  
Db 2392 GATGACGACGATCCTCCAGCTTCTTTGGCGGATCCTGAAAGTTGAATTCACCTGAATGG 2451  
Qy 491 PropheValValGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 2452 CCTATTTTGTGGTGTGTATTTTGTGCTAATAATGAGGCTTGCACACGACGATTC 2511  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAspLysThrThrLeu--- 529  
Db 2512 TCCATAATTTCAAGGTGTAGGGTTTTTACAAAATAATGACACCCCTGAAATCCAG 2571  
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549

Db 2572 CGCAGAACAGCAACTGTTCTTTCTTTATTTGTTCTGATCTCTGGGATCATCTCTTTTCATT 2631  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuLeuThrMetArgLeu 569  
Db 2632 ACGTTTTCCTTCAAGGCTTCACATTTGGCAAGCTGGAGAGATCTTCACCAAGCGACTC 2691  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589  
Db 2692 CGATACATGCTCTTCAATCCATCTGAGACAGGACATAGCTGGTTGTGATGACCCCTAAA 2751  
Qy 590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 2752 AACACCAAGGAGGCTGACACAGGCTTGGCAATGACGCTGCTCAAGTGAAGGGGCT 2811  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 2812 ACAGGGTCTAGGCTTGTGTTATTACCCAGAACATAGCAATCTTGGGACAGGCATCATC 2871  
Qy 630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 2872 ATATCCCTGATCTACGGCTGGCAATTCACACTTTTACTCTAGCAATGTTCCCATCAT 2931  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 2932 GCTATAGCAGGAGTGGTTGAAATGAAATGTTGTCTGACAAAGCGCTGAAAGATAAGAG 2991  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 2992 GAACCTAAGAGTTCTGGGAAGTCTGACAGAACATTTGAAACTTTTCGACACTGTCGTC 3051  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 3052 TCTTTGACTCGGGAGCAGAGTTTGAACATATGATGCCAGAGCTTGCAGATACCATC 3111  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 3112 AGAAATGCTTTGAAGAAAGCGACAGCTTTTGGGATCATCTTTCTCTTCACCCAGGCGCATG 3171  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749  
Db 3172 ATGATTTCTCTATGCTGCTTGTTCGGTTTGTATGCTTACTTGGTGGCAGGAATC 3231  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 3232 ATGACATTTGAAATGTTCTCTAGTATCTCAGCTATGCTTTGGTGGCAGTGCAGTG 3291  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 3292 GGGCAGGTCAGTTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCACATC 3351  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 3352 ATCAGGATCATTTGAAATCCCTGAGATTACAGCTACAGCAGGAGGCTTGAAGCCT 3411  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 829  
Db 3412 AATATGTTGGAGGAAGTGAATTAATGAGTTCATGTTCAACTATCCACCCGACCC 3471  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 3472 AACATCCAGTGTCTCAGGCGCTGAGCCTAGAGGTGAAGAAAGGCGACAGCTGCGCCCTC 3531  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp 869  
Db 3532 GTGGCAGCAGTGGCTGCGGAAGATGACGTGGTCCAGCTGCTGAGCGCTTCTATGAC 3591  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValPheAlaLysGluLeuAsnValGlnTrp 889  
Db 3592 CCCATGGCGGAGACAGTGTCTTAGATGGCAAGAAATAAGCAACTCAATGTCAGTGG 3651  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 3652 CTCGCGCCCACTGGGCATTTGTGTCAGGAGGCCATCTCTGTTGACTGCAGCATCGCC 3711

Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 3712 GAGAACATTCCTACGGAGACAAACAGCCGTCGTCTCTCATAAGGAGATCGTGAAGGCA 3771  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 3772 GCCAAGGAGGCCAACATCCACAGTTTCATCGACTCATCTCCCTGAGAAATACCAACACCA 3831  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 3832 GTGGGAGACAAAGGACTCAGCTGTCGGCGGCGGAGAGCAGCGCATCGCCATCGCGCGC 3891  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsn 989  
Db 3892 GCCCTGTCAGACAGCTCCTCATCTTTACTTCTTGATGAAGCGACATCAGCTCTGGATACG 3951  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 3952 GAGAGTGAAGAGTCTCCAGGAAGCGCTGGACAAAGCCAGGAGGCGCACCTGCATT 4011  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 4012 GTGATCGCGCACCGCTCTCCACCATCCAGACGCGACTTGTCTGTGTGATTCAGAAC 4071  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 4072 GGCAGGTCAAGGAGCAGCGCACCCACAGAGCTCTGCCCCAGAAAGGCGCATCTATTTC 4131  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 4132 TCGATGTCAGTGTGCAGGCT 4152

## RESULT 9

US-09-866-866A-1  
; Sequence 1, Application US/09866866A  
; Patent No. US2002010244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 3860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-866-866A-1

Alignment Scores:  
Pred. No.: 1,73e-280 Length: 3860  
Score: 3100.50 Matches: 595  
Percent Similarity: 75.54% Conservative: 211  
Best Local Similarity: 55.76% Mismatches: 218  
Query Match: 58.17% Indels: 43  
DB: 10 Gaps: 6

US-09-873-409-4 (1-1058) x US-09-866-866A-1 (1-3860)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23  
Db 712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGAGCAGTAGCTGAAGAGTC 771  
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43







1285	AGTGGCTGTGGGAAGACCAACAGTCTCCAGCTGTATGCAGAGGCTCTATGACCCACAGAG	1344
Qy	GlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnValAArgHisTyrArgAsp	263
1345	GGGATGTCAGTGTTCATGGACAGGATATTAGGACCATTAATGTTAAGGTTTCTACGGAA	1404
Qy	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	283
1405	ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTGTTGCCACCACTAGTCTGAAACATT	1464
Qy	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAalaArgGluAlaAsn	303
1465	CGCTATGGCCGTGAATATGTCACCATGGATGAGATTGAGAAGCTGTCAAGAAGCCAAAT	1524
Qy	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	323
1525	GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG	1584
Qy	AlaGlnMetSerGlyGlnLysGlnAArgIleAlaIleAalaArgAlaLeuValArgAsn	343
1585	GCCAGTTGAGTGGTGCGCAGAACAGAGGATCGCCATTTCACCTGGTCCCTCGTTCGCAAC	1644
Qy	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	363
1645	CCCAAGATCTCTCTCTGGATGAGGCCACGTGAGCTTTGGACAAAGGCAAGCAGTGT	1704
Qy	ValGlnAalaLeuGluLysAlaSerLysGlyAArgThrThrIleValValAlaHisAArg	383
1705	GTTCAGGTGGCTCTGGATAAGCCAGAAAGGTTCGACCAACATTGTGATAGCTCATCGT	1764
Qy	LeuSerThrIleAArgSerAlaAepLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	403
1765	TTGTCTACAGTTTCGTAAATGCTGACGTCATCGCTGGTGTTCGATGATGGAGTCATTTGTGGAG	1824
Qy	LysGlyAlaHisAlaGluLeuMetAlaLysAArgGlyLeuTyrTyrSerLeuValMetSer	423
1825	AAAGGAAATCATGATGAACTCATGAAGAAGAAAGCATTTACTTCAAATTTGTCCACAATG	1884
Qy	Gln-----AspIleLysLysAlaAepGluGlnMetGluSerMetThr	437
1885	CAGACACAGCAAAATGAAGTTGAAATGTAGAAATGCAGCTGATGAATCCAAAAGTCAAAAT	1944
Qy	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	452
1945	GATGCCTTGGAATGTCTTCAAAATGATTCAAGATCCAGTCCAGTCTAATAAGAAAAAGATCAACT	2004
Qy	---LysSerIleLys---SerAspPheIleAepLysAlaGluLeuSerThrGlnSerLys	470
2005	CGTAGGAGTCTCGGTGGATCAACAGCCCAAGACAGAAAGCTTTAGTACCAAAAGGGCTCTG	2064
Qy	GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr	490
2065	GATCAAGATATACCTCCAGTTTCCTTTTCGAGGATTTAAGAGCTTAAATTTAACTGTAATGG	2124
Qy	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	510
2125	CCTTATTTTGTGTGTGTATTTCCTTATTAATATAATGGAGGCTCCCAACACGACATTT	2184
Qy	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	529
2185	GCAATAATATTTTCAAAAGATATATAGGGGTTTTTTTACAAGAAATTTAGTATCTGTAACAAAAA	2244
Qy	LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	549
2245	CGACAGAAATAGTAACCTTGTTTTTCACTATTGTTTCTAGCCCTTGGAAATTAATTTCTTTTATT	2304
Qy	SerTyrPheMetGlnGlyLeuPheTyrGlyAArgAlaGlyGluIleLeuThrMetArgLeu	569
2305	ACATTTTTCCTTCAAGGTTTTCATTTTGGCAAGCTGGAGAGATCTCTCACCACAGGGCTC	2364
Qy	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu	589
2365	CGATACATGTTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGTATGACCTTAAA	2424

Qy	590	AnsSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	609
Db	2425	AACCACTCGGAGCATTTGACTACAGGCTCGCAATATGATGTGCTCAAGTTAAAGGGGCT	2484
Qy	610	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	629
Db	2485	ATAGGTTCCAGGCTTGCTGAATATCCAGAAATATAGCAAAATCTTTGGGACAGGAATAATT	2544
Qy	630	IleSerPheIleTyrGlyTyrTpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	649
Db	2545	ATATCCTTCATCTATGGTGGCAACTAACTACTGTACTCTTAGCAATTGTACCCATCATT	2604
Qy	650	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	669
Db	2605	GCAATAGCAGGAGTTGTTGAATATGAAAAATGTTCTGTGCAACAGCACTGAAGATAAGAAA	2666
Qy	670	GluLeuLysHisIleGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	689
Db	2665	GAACTAGAAGGTGCTGGGAAGATCGCTACTCTGAAGCAATAGAAAAATCTCCGAACCGCTTGT	2724
Qy	690	SerLeuThrArgGlnLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	709
Db	2725	TCCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGTCTCAGAGTTTGCAGGTACCATC	2784
Qy	710	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	729
Db	2785	AGAAACTCTTTGAGGAAGCACACATCTTTTGGAAATTACATTTTCTTACCACGAGCAATG	2844
Qy	730	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	749
Db	2845	ATGTAATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCCTACTGTGGTGGCACAATAA	2904
Qy	750	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	769
Db	2905	ATGAGCTTGAGAGATGTTCTGTAGTATTTTTCAGCTGTGTCTTTGGTGCCATGCCCGTG	2964
Qy	770	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	789
Db	2965	GGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAGCCAAAAATATCAGCAGCCACACATC	3024
Qy	790	PheAlaLeuLeuGlnLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	809
Db	3025	ATCATATGATCATTTGAAAAAACCCTTTTGATTGACAGCTACAGCACGGAAGGCCCTAAT	3084
Qy	810	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	829
Db	3085	AACATTTGGAAGAAATGTCACATTTGTTGAAGTTGTATTTCACTATCCACCCGACCG	3144
Qy	830	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	849
Db	3145	GACATCCAGTGCTTCAGGACATGACCTGGAGGTGAAGAGGCCAGACGCTGGCTCTG	3204
Qy	850	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp	869
Db	3205	GTGGGCAGCAGTGGCTGTGGGAAGACACAGTGTGTCCAGCTCCTGGAGCGGTCTTACGAC	3264
Qy	870	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	889
Db	3265	CCCTTTGCGAGGGAAGTGCTGTGTTGATGCCAAGAAATAAAGCCGACTGAATGTTCCAGTGG	3324
Qy	890	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCyserIleAla	909
Db	3325	CTCCAGCACCTGGGCATGTTGTTCCAGGAGGCCCATCTCTGTTTACTGTGACGACATTCGT	3384
Qy	910	GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla	929
Db	3385	GAGAAACATTGCCTATGGAGACACACCCGGTGTGTCTCAGGAAGAGATCGTGAGGCA	3444
Qy	930	AlaAsnAlaIleAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	949
Db	3445	GCAAGGAGGCGCAACATACATGCTTTCATCGACTCAGCTGCCTTAATAAATATAGCAATAA	3504



```
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArg 969
Db 3505 GTAGGACAAAGCACTCAGCTCTGTGGCCAGAAACAAACGCAATGCCATAGCTCGT 3564
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db 3565 GCCCTGTGACACGCTCATATTTGCTTTGGATGAAGCCAGCTCAGCTCTGGATACA 3624
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db 3625 GAAAGTGAAGAGGTGTCCAGAAAGCCCTGGACAAAGCCAGAGAGCCGACCTGCAT 3684
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db 3685 GTGATTGCTCACCCGCTGTCCACCATCCAGATGCGACATTAATAGTGGTTCAGAA 3744
Qy 1030 GlyLysIleLysGlnGlnGlyThrHisGlnGlnLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 3745 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTT 3804
Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 3805 TCAATGGTCAAGTGTCCAGGCT 3825
RESULT 11
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta71m4
US-09-306-417-1
```

```
Alignment Scores:
Pred. No.: 5,71e-280 Length: 8630
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 10 Gaps: 6
US-09-873-409-4 (1-1058) x US-09-306-417-1 (1-8630)
Qy 4 SerLeuThrSerLysGlnLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 1931 TCATTTACTGATAAAGAACTCTTAGCGGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGTC 1990
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuAlaArgSer 43
Db 1991 TTGGCAGCAATTAAGAACTGTGATTGCTTTGGAGGACAAAGAAAGAACTTGAA----- 2044
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 2044 ----- 2044
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysaspPheGlyIleLys 83
Db 2045 -----AGGTACAAACAAAAATTTAGAAAGAGCTAAAGAAATTTGGATAAAG 2089
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 2090 AAGAGCTATTACAGCAATATTTCTATAGGTGGCTGCTTCTCGCTGATCTATGATCTTAT 2149
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 2150 GCTCTGGCCTTCTGTTATGGGACCACCTTGCTCTCAGGGAA-----TATCTATT 2203
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 2204 GGACAGTACTACTGTATTCTTTCTGTTAATTTGGGGCTTTTAGTTGGACAGGCA 2263
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 2264 TCTCCAAGCATTTGAAGCATTTGCAATTCAGAGAGGAGCAGCTTATGAAATCTTCAAGTA 2323
Qy 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 2324 ATTGATAATTAAGCCAGTATTGACAGCTATTCGAGAGTGGGCACAAACCCAGATAATT 2383
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 2384 AAGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCTATCTCGAAAGAAAGTTAAG 2443
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 2444 ATCTTGAAGGGCCTGAACCTGAAGGTGGCAGAGTGGGCGAGCGGTGGCCCTGGTGGAAAC 2503
Qy 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 2504 AGTGGCTGTGGAGAGACACACAGTCCAGCTGATGAGAGGGCTCTATGACCCACAGAG 2563
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 2564 GGGATGCTCAGTGTGATGTCACAGGATATTAGGACCATAAATGTAGAGTTTCTACGGGAA 2623
Qy 264 HistGlyValValSerGlnGluProValIleuPheGlyThrThrIleSerAsnIle 283
Db 2624 ATCATTTGGTGTGGTGAAGAACCTGTATTGTTTCCACCAGTAGCTAGTAANAACATT 2683
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 2684 CGCTATGGCGGTGAAATGTCACCATCGATGAGATTCGAGAAAGCTGTCAAGGAAGCAAT 2743
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 304 ----- 323
```

Db 2744 GCCTATGACCTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGGTTGAGAGAGAGGG 2803  
Qy 324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaLeuValArgAsn 343  
Db 2804 GCCCAGTTGAGTGGGACAGACAGAGATCGCATTCGACGTCCTGGTTGCGAAC 2863  
Qy 344 ProIysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysAla 363  
Db 2864 CCCAAGATCCTCTGCTGGATGAGGCACCTGAGCCTTGACACAGAAAGCGAAGCAGTG 2923  
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 2924 GTTCAGTGCTCTGGATAGGCAGAAAGTCGGACCACTTGTGATAGCTCATCGT 2983  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValIleThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 2984 TTGCTACAGTTGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTGTGGAG 3043  
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer 423  
Db 3044 AAAGGAATCATGATGAACCTCATGAAGAGAAAGGCACTTACTTCAAACTTGTCAAAATG 3103  
Qy 424 Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr 437  
Db 3104 CAGACAGCAGAAATGAAGTTGAATTAGAAATGACGCTGATGAATCCAAAGATGAAT 3163  
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 3164 GATGCCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAGAAAGATCAACT 3223  
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470  
Db 3224 CGTAGGAGTGTCCGTGATCACAGGCCAACAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3283  
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuLysProGluTyr 490  
Db 3284 GATGAAGTATACCTCCAGTTTCTTTGGAGGATTAAGAGCTAAATTTAACTGAATGG 3343  
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 3344 CCTATTCTTGTGTGTATTTTGTGGCCATTATATAATGGAGGCTGCAACCGCATTT 3403  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrLeu 529  
Db 3404 GCAATAATATTTCAAAGATTATAGGGTTTACAAAGATTGATGATCCTGAAACAAA 3463  
Qy 530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 3464 CGACAGAAATAGTAACCTGTTTTCATTATTTCTAGCCCTTGGAAATATTTCTTTATT 3523  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 3524 ACATTTTCTTCAAGTTTTCATATTGGCAAGCTGGAGAGATCTTACCAAGCGGCTC 3583  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleThrPheAspGluLysGlu 589  
Db 3584 CGATACATGTTTCCGATCCATGCTCAGACAGGATGTGATGTTGATGACCCCTAAA 3643  
Qy 590 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 3644 AACACCACTGGACATTCAGTACAGGCTGCCAATGATGCTGCTCAAGTTAAAGGGGCT 3703  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 3704 ATAGTTTCCAGGCTTGTGTAATTTACCAGATATAGCAATCTTGGACAGAGATTAAT 3763  
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 3764 ATATCTTTCATGTTGTCACACTAACACTGTTACTTCTTAGCAATGTGACCATCAT 3823  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 3824 GCAATAGCAGGAGTTGTTGAATGAAATGTTGTCTGGCAAGCACTGAAAGATAAGAAA 3883

Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 3884 GAACTAGAAGAGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTTCCGAACCGTTGTT 3943  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 3944 TCTTTGACTCAGAGCAGGAAGTTTGAACATATGATGTCTCAGAGTTTGCAGGTACCATATC 4003  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 4004 AGAAACTCTTTGAGGAAGCAGACATCTTTTGGAAATACATTTTCTTCCACCGCAATG 4063  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValTyrIleLeuIleGlnAlaGlyArg 749  
Db 4064 ATGATATTTTCTATCTGCTGATGTTTCCGCTTGGAGCCTACTTGTGGGCACATAAATC 4123  
Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 4124 ATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTCTTGTGTCATGGCCGTG 4183  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 4184 GGGCAAGTCAGTTCATTTCTCTGACTATGCCAAAGCCAAATATATCAGCAGCCCATC 4243  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 4244 ATCATGATCATTTGAAAAACCCCTTTGATTCAGCTACAGCAGCAAGGCTTAATGCCG 4303  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValIleSerPhePheTyrProCysArgPro 829  
Db 4304 AACCAATTGGAAGAAATGTCACTTTGTTGTAAGTTGTATTCAACTATCCACCCGACCG 4363  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 4364 GACATCCAGTGTCTCAGGACTGAGCTTGGAGGTGAAGAGGCCAGAGCTGGCTCTG 4423  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 4424 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCTCAGCTCTCTGGAGCGGTTCTACGAC 4483  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889  
Db 4484 CCTTGGCAGGGAAGTGTCTGTGATGTCACAAAGAAATTAAGACGACTGAATGTTCAGTGG 4543  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 4544 CTCGAGACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCACATGCT 4603  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 4604 GAGAACATTCCTATGAGACACACAGCCGGTGTCTCACAGGAAGAGATCGTGAGGCA 4663  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 4664 GCAAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCTTAATAATATAGCACTAAA 4723  
Qy 950 ValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 4724 GTAGGAGACAAAGGAATCAGCTCTCTGTGGCCAGAACCAACGCAATGCGATGCTCGT 4783  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 4784 GCCCTTGTAGACAGCTCATATTTTGTCTTTGATGAAGCCAGCTCAGCTCTGGATACA 4843  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 4844 GAAAGTGAAGAGTTTGTCCAAGAAAGCCCTGGAACAAAGCCAGAGAGCGCCACCTGCAAT 4903  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029  
Db 4904 GTGATTGCTCACCGCCTGCCCATCCAGATGCAGACTTAATAGTGGTGTTCAGAAAT 4963

```
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db 4964 GGCAGAGTCAAGAGCAGTGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5023

Qy 1050 LysLeuValAsnAlaGlnSer 1056
Db 5024 TCAATGGTCAGTGTCACGGCT 5044

RESULT 12
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SPbeta91msA1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: msA1 mdr1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.: 5,71e-280 Length: 8630
Score: 3100.50 Matches: 595
Percent Similarity: 75.54% Conservative: 211
Best Local Similarity: 55.76% Mismatches: 218
Query Match: 58.17% Indels: 43
DB: 10 Gaps: 6

US-09-873-409-4 (1-1058) x US-09-306-417-2 (1-8630)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 23
Db 1931 TCATTACTGATAAAGAACTCTTAGCGTATGCCAAGCTGAGCAGTAGCTGAAGAGTGC 1990
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuArgSer 43
```

```
Db 1991 TTGGCAGCAATTAGAACTGTGATTGGATTGGAGGACAAAAGAAAGAACTTGAA----- 2044
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTppPheTyrPheProGlnTrpLeuLeuSerCys 63
Db 2044 ----- 2044
Qy 64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db 2045 -----AGGTACAACAAAATTTAGAGAAGCTAAAGAAATTGGATAAAG 2089
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db 2090 AAAGCTATTACAGCAATATTTCTATAGGTGCTGCTTTCTCTGCTGATCTATGATCTTAT 2149
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db 2150 GCTCTGGCCTTCGGTATGGACACCTTGGTCTCTCAGGGGAA-----TATTCTATT 2203
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db 2204 GGACAAGTACTCACTGATTCTTTCTGATTAAATTGGGGCTTTTAGTGTGGACAGGCA 2263
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db 2264 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 2323
Qy 164 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db 2324 ATTGATAATAAGCCAAAGTATTGACAGCTATTGGAAGAGTGGGCACCAACAGATAATATT 2383
Qy 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db 2384 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAGTTAAG 2443
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db 2444 ATCTTGAAGGGCTGAACCTGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTTGGAAC 2503
Qy 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db 2504 AGTGGCTGGGGAAGAGCACACAGTCCAGCTGATGACAGGGCTCTATGACCCACAGAG 2563
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db 2564 GGGATGGTCAGTGTTCATGTCAGACAGATATTAGGACCAATATCTAAGTTTCTACGGAA 2623
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 283
Db 2624 ATCATTTGGTGGTGGTGCAGGAACCTGTATTGTTGCCACACGATAGCTGAAAACATT 2683
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db 2684 CGCTATGGCCGTGAAAATGTCACCATGATGATGAGAAAGCTGTCAAGGAAGCCCAT 2743
Qy 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db 2744 GCCTATGACTTTATCATGAACTGCCTCATAAATTTGACACCCCTGGTGGAGAGAGAGGG 2803
Qy 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db 2804 GCCCAGTTGATGGTGGCAGAAAGCAGAGGATCGCCATTGTCAGCTGCGCTGGTTCGCAAC 2863
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db 2864 CCCAAGATCTCTCTGCTGGATGAGGCCACCTGAGCTTGGACACAGAAAGCAGAGTGT 2923
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db 2924 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTGGCAGCCACCATTTGTAGTCTCATCGT 2983
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
```

Db 2984 TTGCTACAGTTTCGTAAATGCTGAGTCATCGCTGGTTTCGATGATGAGTCATTTGGAG 3043  
Qy 404 LygGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 3044 AAAGGAAATCATGATGAATCATGAAAGAAAGGCAATTTACTTCAAACTTGTCACAATG 3103  
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437  
Db 3104 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGCGACTGATGAATCCAAAGTGAAT 3163  
Qy 438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 452  
Db 3164 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACT 3223  
Qy 453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470  
Db 3224 CGTAGGAGTTCGCTGGATCACAAGCCCAAGCAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3283  
Qy 471 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAnLysProGluTyr 490  
Db 3284 GATGAAAGTATACCTCCAGTTTCTTTTGAGGATTAAGAAGCTAAATTTAACTGAATGG 3343  
Qy 491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510  
Db 3344 CCTATTGTTGTTGGTGTATTTTGTGCAATTAATAATGGAGCTCGCAACAGCAATTT 3403  
Qy 511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529  
Db 3404 GCAATAATATTTTCAAGATTATAGGGTTTTCACAGAAATTCATGATCCTGAAACAAAA 3463  
Qy 530 LysHisAspAlaGluLysThrSerMetIlePheValIleLeuGlyValIleCysPheVal 549  
Db 3464 CGACAGAAATAGTAACTTGTTCACATTTCTAGCCCTTGGAAATATTTCTTTTAT 3523  
Qy 550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569  
Db 3524 ACATTTTCTTCAAGTTTTCACATTTGGCAAGCTGGAGATTCCTCAACAGCGGCTC 3583  
Qy 570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 589  
Db 3584 CGATACATGTTTTCGATCCATGCTCAGACAGGATGAGTTGGTTTGATGCCCTAAA 3643  
Qy 590 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609  
Db 3644 AACACCACCTGGAGCATTTGATACAGGCTCGCAATGATGCTCAAGTTAAAGGGCT 3703  
Qy 610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629  
Db 3704 ATAGGTTCCAGGCTTCTGTAAATTTACCAGAAATATAGCAATCTTGGGACAGGAATAT 3763  
Qy 630 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649  
Db 3764 ATATCTTTCATCATGTTGGCAACTAACACTGTTACTCTTAGCAATTTGATCCCATCAT 3823  
Qy 650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669  
Db 3824 GCAATAGCAGGAGTTGTGAAATGTTGCTCGCAGCAGCACTGAAAGATATAAGAAA 3883  
Qy 670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689  
Db 3884 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAAGCAATAGAAAACTTCCGAACCGTTGTT 3943  
Qy 690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709  
Db 3944 TCTTTGACTCAGGACGAGAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 4003  
Qy 710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729  
Db 4004 AGAACTCTTTGAGAAAGCACACATCTTTTGAATTAACATTTTCTTCCACCACCGCAATG 4063  
Qy 730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749  
Db 4064 ATGTATTTTTTCTATGCTGGATGTTTCCGGTTTGGAGCTTACTTTGGTGCCACATAAACTC 4123

Qy 750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769  
Db 4124 ATGAGCTTTGAGGATGTTCTGTAGTATTTCAGCTGTTGTCTTGTGGTCATGGCGGTG 4183  
Qy 770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789  
Db 4184 GGCAGATGCTCATTGCTCTCTGATATGCCAAGGCCAAAATATCAGCAGGCCCATC 4243  
Qy 790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809  
Db 4244 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAAGCCCTAATCCG 4303  
Qy 810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829  
Db 4304 AACACATTGGAAGAAATGTACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 4363  
Qy 830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849  
Db 4364 GACATCCCGTGTCTCAGGACTGAGCTGGAGTGAAGAGGCCAGACGCTGGCTCTG 4423  
Qy 850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869  
Db 4424 GTGGCAGCAGTGGCTGTGGAAAGACACAGTGTCCAGCTCTCGAGCGGTTCTACGAC 4483  
Qy 870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 889  
Db 4484 CCCTTGGCAGGAAAGTGTCTGTGTCGAAAGATAAAGCGACTGAAATGTTTCAGTGG 4543  
Qy 890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909  
Db 4544 CTCGAGACACCTGGGCATCGTGTCCAGAGGCCATCTCTGTTGACTGCAGCATTTGT 4603  
Qy 910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929  
Db 4604 GAGAACTTGGCTATGGAGACACAGCCGGTGTGTACAGGAGAGATCGTGAGGCGCA 4663  
Qy 930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949  
Db 4664 GCAAGGAGGCCAACATACATGCTTTCATCGACTCTGCTAATAAATATAGCACTAAA 4723  
Qy 950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969  
Db 4724 GTAGGAGACAAAGAACTCAGCTCTCTGTGGGCCAGAAACACGCAATTGCCATAGCTCG 4783  
Qy 970 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989  
Db 4784 GCCTTGTAGACAGCTCATATTTTGTGATGAAGCCACGCTCAGCTCTGGATACA 4843  
Qy 990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009  
Db 4844 GAAAGTGAAAGGTTGTCCAAGAACCCCTGGACAAAGCCAGAGAACGCCGCTGCATT 4903  
Qy 1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsn 1029  
Db 4904 GTGATTGCTCACCGCTGTCCACATCCAGACTTAATAGTGGTGTTCAGAAAT 4963  
Qy 1030 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1049  
Db 4964 GGACAGTCAAGGAGCATGGCAGCATCAGCAGCTGTGCGACAGAAAGGCATCTATTTT 5023  
Qy 1050 LysLeuValAsnAlaGlnSer 1056  
Db 5024 TCAATGGTCACTGCCAGGCT 5044

## RESULT 13

US-09-866-866A-7

; Sequence 7, Application US/09866686A

; Patent No. US2002010224A1

; GENERAL INFORMATION:

; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

FILE REFERENCE: 1340-1-021CIP2  
CURRENT APPLICATION NUMBER: US/09/866,866A  
CURRENT FILING DATE: 2001-08-30  
PRIOR APPLICATION NUMBER: 09/584,586  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: PCT/US99/11825  
PRIOR FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: 60/086,988  
PRIOR FILING DATE: 1998-05-28  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 7  
LENGTH: 4788  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-866-866A-7

Alignment Scores:  
Pred. No.: 2,06e-279 Length: 4788  
Score: 3090.50 Matches: 588  
Percent Similarity: 75.12% Conservative: 215  
Best Local Similarity: 55.00% Mismatches: 219  
Query Match: 57.98% Indels: 47  
DB: 1.0 Gaps: 6

US-09-873-409-4 (1-1058) x US-09-866-866A-7 (1-4788)

Qy 4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23  
Db 700 TCAATTTACTGATAAGAACTCCATGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 759  
Qy 24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43  
Db 760 TTACGAGCCATCAGAACTGTGATGCGTTTGGAGGACAAAAGAGAACTTGAA----- 813  
Qy 44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63  
Db 813 ----- 813  
Qy 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 814 -----AGGTACAATAACAACTTGGAAAGCTGAAAGAGCTGGGGATAAG 858  
Qy 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
Db 859 AAAGCTATCAGCGCAACATCTCCATGGGTGCGCTTCTCCTTATCTATGTCATCAT 918  
Qy 104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
Db 919 GCTCTGGCATCTGATATGGGACTTCTTGGTCATCTCCAAAGAA-----TACTCTATT 972  
Qy 124 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 143  
Db 973 GGCAAGTGCTCCTCTCTCTTTCCGTTGTTAAATGGAGCATCTCAGTGTGGACAGCA 1032  
Qy 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 1033 TCTCCAAATATTGAAGCCTTCGCCAATGACGAGGAGCAGCTTATGAAGTCTTCAAAATA 1092  
Qy 164 IleAspLysPheProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 1093 ATTGAATAATAGCCAGATATAGACAGCTTCTCAAGAGTGGGACAAACACACACATA 1152  
Qy 184 GluGlyThrValGluPheLysValSerPheAsnTyrProSerArgProSerIleLys 203  
Db 1153 CAAGGAATCTCGAATTTAAGAATATTCATCTCAGTTACCCATCTCGAAAGAGTTCAG 1212  
Qy 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1213 ATCTTGAAGGGCTCAATCTGAAGGTGAAGCGGACGACGCGTGGCCCTGTTGGCAAC 1272  
Qy 224 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAsp 243  
Db -----

Db 1273 AGTGGCTGTGGAAGAACACAACTGTCCAGCTGATGCAAAAGGCTCTACGACCCCTAGAT 1332  
Qy 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1333 GGCATGTCAGTATCGACGACAGGACATCAGAACCATCATGTGAGGTATCTCAGGGAG 1392  
Qy 264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 283  
Db 1393 ATCATTTGGTGTGTGAGTCAGGAACCTGTCTGTTTCCACACGATCGCGGAGAACATT 1452  
Qy 284 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1453 CGCTATGCGCGAGAAGATGTCCACATGGATGAGATGAGAAAGCTGTCAAGAAAGCCAA 1512  
Qy 304 AlaTyrAspPheIleMetGluPhePheAsnLysPheAsnThrLeuValGlyGluLysGly 323  
Db 1513 GCCTATGACTTCATCAAACTGCCCAACCAATTTGCACCCCTGGTTGGTGAGAGGG 1572  
Qy 324 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1573 GCGCACGTGAGTGGGGACAGAAACAGAAATCGCCATTGCGCGGCCCTGGTCCGCAAT 1632  
Qy 344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1633 CCCAAGATCTCTTTTGTGGACGAGGCCCTCAGCCCTGGATACAGAAAGTGAAGCTGTG 1692  
Qy 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 1693 GTTCAGCGCCGACGTGATAAGGCTAGAGAAAGCGGACCACTTGATGTAGTCTCATCGC 1752  
Qy 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1753 TTGTCTACCGTCTCTAATGCTGACGTCATCTGCTGTTTGTGTTGCTGTTGTTGAG 1812  
Qy 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423  
Db 1813 CAAGCAATCATGATGATGCTCATGAGAGAAAGGCACTTACTTCAAACTTGTCTATGACA 1872  
Qy 424 Gln-----AspIleLysLysAlaAspGluGlnMet 433  
Db 1873 CAGACAGCAAGAAATGAATTTAGAAATAGAGCTTGTAAATCTAAGGATGAATTT 1932  
Qy 434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLys 453  
Db 1933 GATAATTTAGACATGCTCTCAAAAGATTCCAGATCCAGTCTAATAAGAAAGA-----AGA 1986  
Qy 454 SerIleLysSerAspPheIle-----AspLysAlaGluLysSerThrGln 468  
Db 1987 TCAACTCGCAAAAGCATCTGTGGACCATGACCAAGACAGGAGCTTAGTACCAAGAG 2046  
Qy 469 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 488  
Db 2047 GCCTGGATGAAGATGATACCTCCAGTCTCTTTGGCGGATCTCTGAAGTTGAATCAACT 2106  
Qy 489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508  
Db 2107 GAATGGCTTATTTTGGTTGGTATATCTGTGCCATAATAATGGAGGCTTACAGCA 2166  
Qy 509 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThr 527  
Db 2167 GCATTCCTCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGGTGGCCCCCTGAA 2226  
Qy 528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547  
Db 2227 ACCAGCGCGACAGACCAACTGTTTCTGTGTGTTCTGTATCTCTGGATCATTTCT 2286  
Qy 548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567  
Db 2287 TTCATTACATTTTCTTCAGGCTTCATTTGGCAAGCTGGAGAGATCTCTACCAAG 2346  
Qy 568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587  
Db 2347 CGACTCCGATACATGTTTCAAAATCCATGCTGAGACAGGATGTGAGCTGTTTGTATGAC 2406

Qy	948	ThrClnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle	967
Db	3487	ACCAAGATTAGGACAAAGCACTCAGCTGTTCGGGTGGGAGAACGCGCATCGCCATC	3546
Qy	968	AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu	987
Db	3547	GCAGCGCCCTCGTCAGACAGCTCACATTTTACTTCTTGGACCAAGCAACATCAGCTCTG	3606
Qy	988	AspAsnSerGlnLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr	1007
Db	3607	GATACAGAAAGTGAAAGGTTGTCAGGAAGCGCTGGACAAAGCCAGGGAAGCGGCACC	3666
Qy	1008	CysLeuValValThrHisArgIleSerAlaLleGlnAsnAlaAspLeuIleValValLeu	1027
Db	3667	TGCATTTGTGATCGCTCACCGCTGTCCACCATCCAGAACCGGAGCTTATCGTGTGATT	3726
Qy	1028	HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle	1047
Db	3727	CAGAACGGCAAGGTCAAGGAGCAGCGCACCCACAGCAGCTGCTGCGCAGAGGGCATC	3786
Qy	1048	TyrPhelLysLeuValAsnAlaGlnSer	1056
Db	3787	TACTTCTCAATGGTCAGTGTGCAGGCT	3813
RESULT 14			
US-09-880-107-2299			
; Sequence 2299, Application US/09880107			
; Patent No. US20020142981A1			
; GENERAL INFORMATION:			
; APPLICANT: Horne, Darci T.			
; APPLICANT: Vockley, Joseph G.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Gene Logic, Inc.			
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			
; FILE REFERENCE: 44921-5028-WO			
; CURRENT APPLICATION NUMBER: US/09/880,107			
; CURRENT FILING DATE: 2001-06-14			
; PRIOR APPLICATION NUMBER: US 60/211,379			
; PRIOR FILING DATE: 2000-06-14			
; PRIOR APPLICATION NUMBER: US 60/237,054			
; PRIOR FILING DATE: 2000-10-02			
; NUMBER OF SEQ ID NOS: 3950			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 2299			
; LENGTH: 3924			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234			
US-09-880-107-2299			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-873-409-4 (1-1058) x US-09-880-107-2299 (1-3924)			
Qy	4	SerLeuThrSerLysGlnLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal	23
Db	750	GCATTTAGTGACAAAGCACTAGCTGTTATGCAAAAGCAGGCCCTGGCAGAGAGCT	809
Qy	24	LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer	43
Db	810	CTGGGGCCATCAGGACTGTGTAGCTTTTCGGGGCCAGAACAAAGACCTGAA-----	863
Qy	44	PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuSerCys	63
Db	863	-----	863



QY 64 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83  
Db 864 -----AGGTATCAGAAACATTTAGAAAATGCCAAAGAGATTGGAATTTAAA 908  
QY 84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103  
Db 909 AAAGCTATTTCAGCAAAACATTTCCATGGGTATTCCCTTCTGTTTAATATATGATCATAT 968  
QY 104 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123  
Db 969 GCATGGCCCTTCTGGTATGGATCCACTCTAGTCATATCAAAAGAA-----TATACATATT 1022  
QY 124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143  
Db 1023 GGAATGCAATGACAGTTTTTTTTTCAATCCTAATTGGAGCTTTTCAGTGTGGCAGGCT 1082  
QY 144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163  
Db 1083 GCCCATGATTGATGCTTTTGCATATGCAAGAGGAGCAGCATATGATCTTTTGATATT 1142  
QY 164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183  
Db 1143 ATTCATTAATAATCTAAATATGACAGTTTTTTCAGAGAGAGACAAACACGACAGCATC 1202  
QY 184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203  
Db 1203 AAAGGGAATTTGGAGTTCAATGATGTTTCACTTTTCTTACCTTCTCGAGCTAACGTCAG 1262  
QY 204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223  
Db 1263 ATCTTGAAGGGCCCTCAACCTGAAGGTGCAGAGTGGGAGAGCGGTGGCCCTGGTTGGAGT 1322  
QY 224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243  
Db 1323 AGTGGCTGTGGGAAGAGACAAACGGTCCAGCTGATACAGAGCTCTATGACCCCTGATGAG 1382  
QY 244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263  
Db 1383 GGCACAAATTAACATTGATGGCCAGGATATTAGGAACTTTAAATGTAACATATCTGAGGGA 1442  
QY 264 HisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle 283  
Db 1443 ATCATTTGGTGTGGTCAGTCAGAGCCGGTGTGTTTTCCACCACAAATGCTGGAATAATT 1502  
QY 284 LysTyrGlyArgAspAspValTyrAspGluMetGluArgAlaAlaArgGluAlaAsn 303  
Db 1503 TGTATTGGCCCTGGAATGTAAACCATGATGAGATAAGAAAGCTGTCAAAAGAGGCCAAC 1562  
QY 304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 323  
Db 1563 GCCTATGATTTATCATGAATTTACCACAGAAATTTGACACCCCTGGTTGGAGAGAGAGG 1622  
QY 324 AlaGlnMetSerGlyGlyGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343  
Db 1623 GCCCAGCTGAGTGTGGGAGAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1682  
QY 344 ProlLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363  
Db 1683 CCCAAGATCCCTTCTGCTGGATGAGGCCACGTCAGCATTTGGACAGAAAGGAGAGCTGAG 1742  
QY 364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383  
Db 1743 GTACAGGAGCTCTGGATTAAGGCCAGAGAGCCGGACCAACCATTTGTATGACACACCGA 1802  
QY 384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403  
Db 1803 CTGTCTACGGTCCGAAATGTCAGATGTCATCGCTGGGTTTGGAGTGGAGTAATTGTGGAG 1862  
QY 404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 421  
Db 1863 CAAGGAAGCCACAGCGCACTGATGAAGAAAGGAGGGGTACTTCAAACTTTGTCAACATG 1922  
QY 422 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 434

Db 1923 CAGACATCAGGAAGCCAGATCCAGTCAGAAAGATTGAACTAAATGATGAAAGGCTGCC 1982  
QY 435 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 454  
Db 1983 ACTAGAAATGCCCCCAATAGGCTGGAATATCGCCTATTATTAGGCAATCTACTCAGAAAAAC 2042  
QY 455 IleLys---SerAspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIle 472  
Db 2043 CTTAAAAATTCACAAATGTGTCAAGAGAGCCTTGATGTGGAACCGATGGACTTTGAAGCA 2102  
QY 473 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPhe 492  
Db 2103 AATGTGCCACAGTGTCTTCTTCTGAAGGTCTGAAATGAAATGAAATGAAATGAAATGAAATG 2162  
QY 493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512  
Db 2163 TTTGTCTGTGGAAACAGTATGTGCATTTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTC 2222  
QY 513 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAsp 532  
Db 2223 ATATTCTCAGATCATAGCGATTTTGGACAGGCGATGATGCAGTGAAGCAGCAGAG 2282  
QY 533 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 552  
Db 2283 TGCAACATATTCTTCTTGAATTTCTTGAATTTCTGGGAATATTCTTTTACTTTCTTC 2342  
QY 553 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 572  
Db 2343 CTTTCAGGGTTTCAGTTTGGGAAGCTGGCGAGATCTCCACAGAGAGCTGCGGTCAATG 2402  
QY 573 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThr 592  
Db 2403 GCITTTAAAGCAATGCTTAAGCAGGACATGAGCTGGTTTGTATGACCATAAAAAACAGTACT 2462  
QY 593 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 612  
Db 2463 GGTGCACCTTCTACAGACTTGCACAGATGCTGCCCAAGTCCCAAGGAGGCCACAGGAACC 2522  
QY 613 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 632  
Db 2523 AGGTGGCTTTTAAATGTCACAGATATAGTAACTTGGAACTGGTATTTATCATATCATTT 2582  
QY 633 IleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 652  
Db 2583 ATCTACGGTTGGCAGTTAAACCTATTGCTATTAGCAGTGTTCATATTGCTGTGTC 2642  
QY 653 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 672  
Db 2643 GGAATTTGAAATGAAATTTGTTGGCTGGAATGCCAAAGAGATAAAAAAGAACTGGAA 2702  
QY 673 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 692  
Db 2703 GCTCTCGAAAGATTTGCAACAGAGGCAATAGAAAATATTAGGACAGTGTGTCTTTGACC 2762  
QY 693 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 712  
Db 2763 CAGAAAGAAAATTTGAATCAATGATGTTGAAAAATTTGATGGACCTTACAGAAATCT 2822  
QY 713 SerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 732  
Db 2823 GTGCAGAAAGGCACACATCTATGGAATTTACTTTTAGTATCTCACAGGCAATTTATGTATTT 2882  
QY 733 AlaTyrAlaAlaGlyPheArgPheGlyValTyrIleLeuIleGlnAlaGlyArgMetThrPro 752  
Db 2883 TCCATATGCCGGTGTGTTTTCGATTTGGTGCAATTCATCTGTAATGGACATATGCGCTTC 2942  
QY 753 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 772  
Db 2943 AGAGATGTTATCTGGTGTGTTTCTGCAATTTGTAATTTGGTGGCTCTAGGACATGCC 3002  
QY 773 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 792



Db 3003 AGTTCATTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATCTG 3062  
QY LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 812  
Db 3063 TTTGAAGACAACCTCTCTGATTCAGCTACAGCTGAAGAGGGCTGAAGCTGATAAATTT 3122  
QY 813 GluGlyAsnLeuGluPheArgGluValSerPhePheThrProCysArgProAspValPhe 832  
Db 3123 GAGGAATAATATAACATTTAATGAAGTCTGTTCAACTATCCACCAGCAACAGTGCCCA 3182  
QY 833 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 852  
Db 3183 GTGCTTCAGGGCTGAGCTCGAGGTGAAGAGCAAGCCAGACACTAGCCCTGGTGGGAGC 3242  
QY 853 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuThrAspProValGln 872  
Db 3243 AGTGGCTGTGGGAAGACACGGTGGTCCAGCTCTCGAGGGTCTTACGACCCCTTGGCG 3302  
QY 873 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 892  
Db 3303 GGGACAGTGTCTTCGATGTGTCAGAGCAAGAAAGAACTCAATGTCCAGTGGCTCAGAGCT 3362  
QY 893 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 912  
Db 3363 CAATCCGAATCGTGTCTCAGAGCCCTATCTATTCAGTCGACATTCGCGAGAAATTT 3422  
QY 913 AlaThrGlyAspAsnSerArgValValProLeuAspGluLysGluAlaAlaAsnAla 932  
Db 3423 GCCTATGGAGACAACACCGGTTGTATCACAGGATGAATTTGTGAGTCAGCCCAAGCT 3482  
QY 933 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysThrAsnThrGlnValGlyLeu 952  
Db 3483 GCCAACATACATCTTTTCATCGAGAGCTTACCCCAAAATATGAACAAGAGTGGGAGAT 3542  
QY 953 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 972  
Db 3543 AGGGGACTCAGCTCTCAGAGGTCAAAACAGAGGATGCTATGCCAGCCCTCATC 3602  
QY 973 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 992  
Db 3603 AGACAACCTCAATCTCTCTGATGAGCTACATCAGCTCTGGATACCTGAAAGTAA 3662  
QY 993 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1012  
Db 3663 AAGGTTCTCAAGAGCCCTCGCAAAAGCCAGAGCCGACCTGCATTTGTGATTGCT 3722  
QY 1013 HisArgLeuSerAlaIleGlnAsnAspLeuIleValValLeuHisAsnGlyLysIle 1032  
Db 3723 CACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGAATGGGAGATC 3782  
QY 1033 LysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleThrPheLysLeuVal 1052  
Db 3783 AAGGAGATGGCAGCATCAGCAGCTCTGTCGACAGAAAGCATCTATTTTCAATGTC 3842  
QY 1053 AsnAlaGlnSer 1056  
Db 3843 AGTGCTCCAGCT 3854

## RESULT 15

US-09-866-866A-5  
; Sequence 5, Application US/09866866A  
; Patent No. US2002010224A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 4189  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-866-866A-5  
Alignment Scores:  
Pred. No.: 5,53e-279 Length: 4189  
Score: 3085.00 Matches: 595  
Percent Similarity: 74.93% Conservative: 206  
Best Local Similarity: 55.66% Mismatches: 222  
Query Match: 57.88% Indels: 46  
DB: 10 Gaps: 7  
US-09-873-409-4 (1-1058) x US-09-866-866A-5 (1-4189)  
QY 1 MetValIleSerLeuThrSerLysGluLeuSerAlaThrSerLysAlaGlyAlaValAla 20  
Db 700 GTATTGACTTTCATTTACTAATAAGAACTCCAGCTTATCCAAAAGCTGGAGCAGTTGCT 759  
QY 21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40  
Db 760 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCCCTTTGGAGGACAAACAGAAGAACTT 819  
QY 41 GlnArgSerPheLeuLeuAsnIleThrArgThrAlaThrPheThrProGlnTrpLeu 60  
Db 820 GAA----- 822  
QY 61 LeuSerCysValLeu\*\*\*PheValArgThrGlnAsnLeuLysAspAlaLysAspPhe 80  
Db 823 -----AGTACATAAATAAATTTAGAGAAGCTAAAAATGTT 858  
QY 81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValThrPhePheMetAsn 100  
Db 859 GGCATAAAGAAAGCTATCACAGCAGCATTTTCATAGGCATTCCTACCTGTTGGTCTAT 918  
QY 101 GlyThrThrGlyLeuAlaPheThrThrGlyThrSerLeuIleLeuAsnGlyGluProGly 120  
Db 919 GCATCATATGCATCGCATTCGTATGGGACATCCTTGGTCTCTCAATATGAA----- 972  
QY 121 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerThrCysIle 140  
Db 973 TATCTATTGGAGAGTCTTACTGCTCTCTCTCTATTTTGTGGGACTTTTAGTATT 1032  
QY 141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160  
Db 1033 GGACACTTGGCCCAACATAGAAGCCTTTGCAACGCGACGAGGGCAGCCTTTGAAATC 1092  
QY 161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyThrLysPro 180  
Db 1093 TTCAGATAATATGAACAGCCAGCATTTGACAGCTTCTCAACAAAGGGCTACAAACCA 1152  
QY 181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnThrProSerArgPro 200  
Db 1153 GACAGTATATGGAACTTAGAGTTTAAATAATGTTCACTTCACTACCCATCGAGAGC 1212  
QY 201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220  
Db 1213 GAAGTTCCAGATCTTGAAGGGCCTCAATCTCAAGGTGAAGAGCGGACAGCGGTGGCTTG 1272  
QY 221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuThrAsp 240  
Db 1273 GTTGGCAACAGTGGCTGTGGAAAAGCACAACTGCTCCAGCTGATGCGAGAGGCTCTACGAC 1332  
QY 241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260  
Db 1333 CCCCTGGAGGGCGGTGTCAGTATCGCGCAAGACATCAGACCATCATGTGAGGTAT 1392  
QY 261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280



```
QY 986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
    ||||| :|||
Db 3607 GCTCTGGATACAGAAAGTGAAGTTGTCCAGGAAGCGCTGGACAAAGCCAGGAAGGC 3666
    ||||| :|||

QY 1006 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
    ||||| :|||
Db 3667 CGCACCTGCATTGTGATCGTCAACCGCTGTCCACCATCCAGAACGCGGACTTGATCGTG 3726
    ||||| :|||

QY 1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1045
    ||||| :|||
Db 3727 GTGATTGAGAACGGCAAGTCAGGAGCACGGCACCCACCAGCAGCTGCTGGCGCAGAG 3786
    ||||| :|||

QY 1046 AspIleTyrPheLysLeuValAsnAla 1054
    ||||| :|||
Db 3787 GGCATCTACTTCTCAATGGTCCAGGCT 3813
    ||||| :|||
```

Search completed: March 31, 2003, 16:04:42  
Job time : 638.104 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 444.798 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-5  
Perfect score: 6159  
Sequence: 1 MILGTLASLVNGACPLMPL.....QELLNRDIYFKLVNAQSVQ 1222

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-O=/cgn2\_1/USPTO pool/US09873409/runat\_27032003\_115424\_19290/app query fasta\_1.7544  
-DB=Published Applications NA -QWTF=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat\_27032003\_115424\_19290  
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/prodata/2/pubpna/PCTU5\_PUBCOMB.seq:
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:
- 10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:
- 11: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:
- 12: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:
- 13: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:
- 14: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6157	100.0	3702	10	US-09-873-409-13
2	5988.5	97.2	3621	10	US-09-873-409-14
3	5328	86.5	3177	10	US-09-873-409-12
4	4561.5	74.1	2856	10	US-09-873-409-10

5	3530.5	57.3	3860	10	US-09-866-866A-1	Sequence 1, Appli
6	3525.5	57.2	4317	9	US-10-044-671-1	Sequence 1, Appli
7	3521.5	57.2	3860	10	US-09-866-866A-3	Sequence 3, Appli
8	3521.5	57.2	8630	10	US-09-306-417-1	Sequence 1, Appli
9	3521.5	57.2	8630	10	US-09-306-417-2	Sequence 2, Appli
10	3517	57.1	4425	10	US-09-769-097-3	Sequence 3, Appli
11	3515	57.1	4369	10	US-09-769-097-1	Sequence 1, Appli
12	3514	57.1	4643	9	US-10-072-621-2	Sequence 2, Appli
13	3503	56.9	4189	10	US-09-866-866A-5	Sequence 5, Appli
14	3499	56.8	4788	10	US-09-866-866A-7	Sequence 7, Appli
15	3481	56.5	3924	10	US-09-880-107-2299	Sequence 2299, Ap
16	3456	56.1	2066	10	US-09-873-409-9	Sequence 9, Appli
17	3453	56.1	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
18	3405	55.3	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	2811.5	45.6	5010	10	US-09-917-800A-483	Sequence 483, App
20	2697	43.8	2021	10	US-09-873-409-15	Sequence 15, Appl
21	2528.5	41.1	1940	10	US-09-873-409-16	Sequence 16, Appl
22	2358	38.3	4175	10	US-09-749-340-3	Sequence 3, Appli
23	2328	37.8	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
24	2259	36.7	4002	9	US-09-758-828-1	Sequence 1, Appli
25	2259	36.7	4002	9	US-09-758-828-3	Sequence 3, Appli
26	2129.5	34.6	3861	9	US-09-938-842A-263	Sequence 263, App
27	2039.5	33.1	3512	10	US-09-749-340-7	Sequence 7, Appli
28	1987	32.3	4653	9	US-10-101-388-2	Sequence 2, Appli
29	1636	26.6	3792	9	US-09-882-694-10	Sequence 10, Appl
30	1617	26.3	2698	10	US-09-749-340-5	Sequence 5, Appli
31	1538.5	25.0	2681	10	US-09-749-340-8	Sequence 8, Appli
32	1526.5	24.8	8036	9	US-10-101-388-1	Sequence 1, Appli
33	1501.5	24.4	3999	9	US-09-882-694-9	Sequence 9, Appli
34	1101.5	17.9	1175	10	US-09-873-409-11	Sequence 11, Appl
35	1080	17.5	1810	10	US-09-749-340-4	Sequence 4, Appli
36	1066.5	17.3	7296	10	US-09-070-927A-59	Sequence 59, Appl
37	1027	16.7	3741	9	US-09-738-626-1251	Sequence 1251, Ap
38	1008.5	16.4	9047	10	US-09-070-927A-12	Sequence 12, Appl
39	981.5	15.9	12438	10	US-09-070-927A-173	Sequence 173, App
40	930.5	15.1	640681	10	US-09-700-988-1	Sequence 1, Appli
41	902.5	14.7	3601	10	US-09-070-927A-255	Sequence 255, App
42	901.5	14.6	6415	10	US-09-070-927A-275	Sequence 275, App
43	770.5	13.5	2298	9	US-10-156-239-18	Sequence 18, Appl
44	770.5	12.5	2298	10	US-09-795-693-18	Sequence 18, Appl
45	770.5	12.5	3408	9	US-10-156-239-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-09-873-409-13  
; Sequence 13, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 3702  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (723)..(723)  
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)  
US-09-873-409-13

Alignment Scores:  
Pred. No.: 0  
Score: 6157.00  
Percent Similarity: 100.00%  
Length: 3702  
Matches: 1222  
Conservative: 0

Best Local Similarity: 100.00%  
Query Match: 99.97%  
DB: 10  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-873-409-5 (1-1222) x US-09-873-409-13 (1-3702)

Qy 1 MetileuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 34 ATGATCTGGGTATACCTGGCATCTGCTCAATGGAGCGCTTCTTCTTAATGCCACTG 93  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 94 GTTTTAGCAAAATAGTATGATTAACCTTATTAGTGGATGTCTAGTCCAAACCTAACACATAC 153  
Qy 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaIleLeuPheGly 60  
Db 154 TCTTTCTTCAGGTGGACCCGTATATAGTTGGAAATAGGTGTGCTGCTTGAATTTTGGT 213  
Qy 61 TyrlleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
Db 214 TACATACAGATTTCTTGTGGATTAATTAACCTGCAGCACAGACAGAGAGATTCGAAAA 273  
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100  
Db 274 CAGTTTTTTTCATTCAGTTTTGGCACAGACATCGGCTGGTTGATAGCTGTGACATCGGT 333  
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTGTATGAGATAAGATT 393  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 394 GCTCTGTTGTTCCAAACATGTCTACTTTTTTCGATGGCTTGGCAGTTGGTTGGTGAAG 453  
Qy 141 GlyTyrPheLysLeuThrLeuValLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 454 GGCTGGAAACTCACCCTAGTGACTCTATCCAGCTCTCTCTTATATAGTGTCTCAGCGCA 513  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 514 GCATGTTCTTAGGATGGTCATCTCATTCAGCAGTAAGGAATTAAGTGCTTATCCAAAGCT 573  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 574 GGGCTGTGGCAGAGAGAGCTTGTGTCATCAATCGAAGACATCATAGCTTTTAGGGCCAC 633  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaThrPheTyrPhe 220  
Db 634 GAGAAAGAACTTCAAGGCTCTTCTCTTTTAAATATATAACAAGATATGCTTGGTTTATTTT 693  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 694 CCCAGTGGCTACTAAGTGTGTCTTCTTNTTGTAGAGTATACAGAAATCTCAAGAT 753  
Qy 241 AlalysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 754 GCAAGGATTTTGGCATAAAAGACATATAGCTTCAAAAGTGCTCTCTGTGTGTGTAC 813  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsn 280  
Db 814 TTTCTTTATGAATGGAACCTATGGAATTTGCTTTTGGTATGGAACCTCTTGTATTTAAT 873  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 874 GGAAGAACTTGATATACATCGGAGCTGTCTTGTCTGTTTCTTTAGTGAATCCATAGC 933  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 934 AGTATTGCAATGGAGCAGCAGCTCCCTCACTTTGAAACCTTCCCAATAGCCCGAGAGCT 993  
Qy 321 AlaPheHisIlePheGlnValIleAspLysProSerIleAspAsnPheSerThrAla 340  
Db 994 GCCTTTTCATATTTCCAGGTATTATGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1054 GGATATAAACCTGATCCATAGAGGAACCTGGGAATTTTAAAAATCTTTCTTCAATTAT 1113  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1114 CCATCAAGACCATCTATCAAGATTTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173  
Qy 381 ThrValAlaIleuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400  
Db 1174 ACAGTCCGCTTGGTCTCAATGGCAGTGGGAAGACGTAGTCAGCTTCTTCGAC 1233  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1234 AGTTATATATGATCCGAGATGATGGCTTTATCATGGTGATGAGATCATCATCAGAGTTTA 1293  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1294 AATGTCCGCAATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCTGTGTTGTCGGG 1353  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1354 ACCACCATCAGTAACAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAG 1413  
Qy 461 ArgAlaAlaArgGluAlaAlaAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1414 AGACGACGACGAGGACCAATGCGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1474 ACATTTGGTAGGGAAAAAGAGCTCAAAATGAGTGGAGGCGCAGAAACAGAGGATCGCAAT 1533  
Qy 501 AlaArgAlaIleuValArgAsnProLysIleLeuIleuAspGluAlaThrSerAlaLeu 520  
Db 1534 GCTCGTGGCTTAGTTGCAAAACCCCAAGATTTCTGATTTTAGATGAGGCTAGCTGTCGCTG 1593  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1594 GATTCAGAAAGCAAGTCACTGTTCAAGCTGCTGAGAGGCGAGCAAGGTCGAGCT 1653  
Qy 541 ThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1654 ACAATCGTGGTAGCACACCGACTTCTTCTACTATTTCGAAGTGCAGATTTGATTTGACCTA 1713  
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1714 AAGGATGGAAATGCTGGCGAGAAAGGACACATGCTGAACTAATGGCAAAACGAGGCTTA 1773  
Qy 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600  
Db 1774 TATTATTCCTTGTGTATGTCACAGGATATTTAAAAAGCTGTGAACAGATGGATCAATG 1833  
Qy 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620  
Db 1834 ACATATTTCTACTGAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGACATCAAG 1893  
Qy 621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640  
Db 1894 TCAGATTTCAATGACAAAGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTTCTCTGAA 1953  
Qy 641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGly 660  
Db 1954 GTCTCTCTATTAATAAATTTTAAAGTTAAACAGCTGAATGGCTTTTGTGGTCTTGGGG 2013  
Qy 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680  
Db 2014 ACATTTGGCTTCTGTTCTAAATGGAACCTGTTTCAATCCAGTATTTTCCATCATCTTTGCAAAA 2073  
Qy 681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700  
Db 2074 ATTATAACCATGTTTGGAAATTAATGATAAAACCAATTAAGCATGATGAGAAATTTAT 2133

```
QY 701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
DB 2134 TCCATGATATTCGTCTATTTGGGTGTTATTTGCTTTGTCAGTTATTTATTCATGCGAGGATTA 2193
QY 721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
DB 2194 TTTTACGGCAGACGAGGGGAAATTTTAAACGATGAGATTAAAGACACTTGGCCCTTCAAGGCC 2253
QY 741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
DB 2254 ATGTTATATCAGGATATTCGCTGTTTGTATGAAAGGAAACACAGCAGAGGCTTGACA 2313
QY 761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
DB 2314 ACAATATTAGCCATAGATATAGCACAAATTTCAAGGAGCAACAGGTTCCAGGATTCGCGTC 2373
QY 781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 800
DB 2374 TTAACACAAATGCAACTAACATGGGACCTTCAGTTATCATTTCTTTATATATGATGG 2433
QY 801 GluMetThrPheLeuIleLeuSerIleAlaProValIleAlaValThrGlyMetIleGlu 820
DB 2434 GAGATGACATTCCTGATTCGTAGTATTCCTCAGTACTTGCCTGACAGGAATGATTGAA 2493
QY 821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
DB 2494 ACCGACGAATGACTGGATTTGCGCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553
QY 841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
DB 2554 ATAGCAACTGAAAGCTTTGGAGAAATATACGTACTATATAGTGTCTATTAACAGGGGAAAGGCC 2613
QY 861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
DB 2614 TTCGACCAATGTATGAGAGATGCTTTCAGACTCAACACAGAAATACCTCGAAGAAGCA 2673
QY 881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
DB 2674 CAGATTATTGGAAGCTGTTATGCTATTCAGCCATGCTTTATATATTTGCTATGCAGCA 2733
QY 901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
DB 2734 GGGTTTCGATTTGGAGCCCTATTTAAATCAAGCTGGACGAATGACCCACAGAGGCATGTTTC 2793
QY 921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
DB 2794 ATAGTTTTTACTGCATTTGCATATGAGCTATGCGCATCGGAAGAAAGCTCGTTTGGCT 2853
QY 941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
DB 2854 CCTGAATATTCCAAAGCCAAATCGGGGCTGCGCATCTGTTGCTTTGTTGGAAGAAAGAA 2913
QY 961 ProAsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeu 980
DB 2914 CCAAAATATAGACAGCCGAGTCAAGAAGGGAAGGAAAGCCAGACATGTGAAGGGAATTTA 2973
QY 981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
DB 2974 GAGTTTCGAGAAGTCTCTTCTCTATTCATGCGCCAGATGTTTTCATCTCCGTGGC 3033
QY 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
DB 3034 TTATCCCTCAGTATTGCGGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGCTGGG 3093
QY 1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
DB 3094 AAAAGCACTCTCTTCAACTTCTGCAGAGACTTTATGACCCCGTGAAGGACAAAGTGTG 3153
QY 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
DB 3154 TTTGATGGGTGGATGCAAAAGAAATTTGAATGTACAGTGGCTCCGTTCCCAAAATAGCAATC 3213
QY 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
```

```
DB 3214 GTTCTCTAAGAGCCCTGTGCTCTTCAACTGCAGCATTGTGTGAGAACATCGCCCTATGGTGAC 3273
QY 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
DB 3274 AACAGCCGTGTGGTCCATTAGATGATCAAGAAGCGCAATGCAGCAATATCCAT 3333
QY 1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
DB 3334 TCTTTTATTGAAGGCTCTCCCTGAGAAATACAAACACACAAGTTGGACTGAAAGGAGCACAG 3393
QY 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaAArgAlaLeuLeuGlnLysProLys 1140
DB 3394 CTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAA 3453
QY 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
DB 3454 ATTTTATTGTTGATGAGCCACTTCAGCCCTCGATATGATGACAGTGAAGAGTGGTTCAG 3513
QY 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
DB 3514 CATGCCCTTGATAAAGCCAGGACGCGGAGGACATGCTAGTGGTCACTCACAGGCTCTCT 3573
QY 1181 AlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGluGlnGly 1200
DB 3574 GCAATTCAGAACCGCATTTGATAGTGGTCTTGCACAATGGAAGATATAAGGAACAAGGA 3633
QY 1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
DB 3634 ACTCATCAAGAGCTCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693
QY 1221 ValGln 1222
DB 3694 GTGCAG 3699
```

## RESULT 2

```
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14
```

## Alignment Scores:

```
Pred. No.: 0 Length: 3621
Score: 5988.50 Matches: 1195
Percent Similarity: 97.79% Conservative: 0
Best Local Similarity: 97.79% Mismatches: 0
Query Match: 97.23% Indels: 27
DB: 10 Gaps: 1
```

US-09-873-409-5 (1-1222) x US-09-873-409-14 (1-3621)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCCTGGGTATATCTGGCATCTGCTCACTGGTCAATGGAGCTGCTCTCTTTAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGAGTGAATAACTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153
```



QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyValAlaAlaLeuIlePheGly 60  
DB 154 TCITTCCTCAGGTTGACCTGTATTATGTTGGATAGGTCTGCTGCCITGATTTTGGT 213  
QY 61 TyrIleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrIleArgLys 80  
DB 214 TACATACAGATTTCTCTGTGGATTATTAACCTGCAGCACGACAGACCAAGAGGATTCGAAA 273  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
DB 274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGCGTGGTTGATAGCTGTGACATCGGT 333  
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTCATGGTATTCGAGATTAAGATT 393  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 394 GCTCTGTGTTGTTCAAAACATGCTACTTTTTCGATTGGCCTGGCAGTTGTTGGTGAAG 453  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 454 GGTGGAAACTCACCTTAGTACTCTATCCACGCTCTCTCTTATAATGGCTTCAGCGGCA 513  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 514 GCATGTTCTAGGATGGTCATCTCATTCACCAAGTAAGATTAAGTGCCTATTCGAAGCT 573  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 574 GGGCTGTGGCAGAGAAGCTTGTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAG 633  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 634 GAGAAAGAACTTCAA----- 648  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 649 -----AGGTATACACGAATCTCAAGAT 672  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 673 GCAAGAGATTTGGCATAAAGAGACTATAGCTTCAAAAGTGCTCTCTGGTGTGTGATC 732  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
DB 733 TTCCTTTATGAATGGAACTATGGAACCTTGTCTTTGGTATGGAACCTCTTGATTCCTAAT 792  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 793 GGAGAACTCGATATACATCGGACCTGTTCTTGCTGTTTCTTTAGTGTAAATCCATAGC 852  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 853 AGTTATTGCAATGGAGCAGCAGCTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912  
QY 321 AlaPheHisIlePheGlnValIleAspLysIleProSerIleAspAsnPheSerThrAla 340  
DB 913 GCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTAGATAGATACTTTTCCACAGCT 972  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 973 GGATATAAACTGNAATCCATAGAAGAACTGTGGAAATTTAAAGTTGTTCTTCAATAT 1032  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1033 CCATCAAGACCATCTATCAAGATTTCTCAAGGCTCTGAATCTCGAAATTAAGTCTGGAGAG 1092  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1093 ACAGTCCCTTGGTGGCTCTCAATGGCAGTGGGAAGATAGCGTAGTCCAGCTTCTGCAG 1152  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420

DB 1153 AGGTTATATGATCCGATGATGGCTTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1213 AATGTGGCGCATATTTCGAGACCATATTGGAGTGGTTAGTCAAGAGCGCTGTTTGTTCGGG 1272  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 1273 ACCACCATCAGTAACAAATATCAAGTATGACCGAGATGATGTCATGATGACAGAGATGGAG 1332  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1333 AGACGACGAAAGGAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1392  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1393 ACATTTGGTAGGGGAAAAGAGCTCAAAATGATGGAGGCGCAAAAACAGAGGATTCGCAATT 1452  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 1453 GCTCTGTCCTTAGTTGCAAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTGCGCTG 1512  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
DB 1513 GATTCAGAAAGCAAGTCAAGCTGTTCAAGCTGCACCTGGAGAGCGCAAAAGGTGGACT 1572  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
DB 1573 ACAATCGTGTAGCACACCGACTTTCTACTATTTCGAAGTGCAGATTTGATTTGTGACCTTA 1632  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
DB 1633 AAGGATGGAAATGCTGGCGGAGAAAGGAGCACATGCTGAACCTAATGGCAAAACGAGGTCTA 1692  
QY 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600  
DB 1693 TATTATTCACTTGATGTGCAGGATATTAATAAAGCTGATGACAGATGCAATG 1752  
QY 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620  
DB 1753 ACATATTCTTACTGAAAGAAAGACCAACTCACTCTCTCTGCTCTGCTGAGAGCATCAAG 1812  
QY 621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640  
DB 1813 TCAGACTTCATTGACAGGCTGAGGAAATCCACCAATCTTAAGAGATTAAGTCTTCTGAA 1872  
QY 641 ValSerLeuLysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGly 660  
DB 1873 GTCTCTCTATTAATAATTTTAAAGTTAAACAAAGCCTGAATGGCTTTTGTGCTTCTGGGG 1932  
QY 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680  
DB 1933 ACATTTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 1992  
QY 681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700  
DB 1993 ATTATAACCATGTTTGGAAATTAATGATAAAACCAATTAAGCATGATGACAGAAATTTAT 2052  
QY 701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720  
DB 2053 TCCATGATATTCGTCATTTGGTGTATTGCTTTGTGCTTATTTTCATGCGAGGATTA 2112  
QY 721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740  
DB 2113 TTTTACGGCAGACGAGGGGAAATTTTAAACGATGAGATTAAACACTTGGCCTTTCAAGGCC 2172  
QY 741 MetLeuTyrGluAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760  
DB 2173 ATGTTATATCAGGATATTCCTGCTTGTGTTGATGAAGAAACAAACGACAGAGGCTTGACA 2232  
QY 761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780

Db 2233 ACAATATTAGCCATAGATATAGCACAAATTCAGGAGCAACAGGTTCCAGGATTGGGCTC 2292  
Qy LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 800  
Db 2293 TTAACACAAATGCAATACATGGGACTTTTCAGTTATCATTTCTTTATATATGATGG 2352  
Qy GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820  
Db 2353 GAGATGACATTCCTGATCTTGAGTATTGCTCCAGTACTTCCCGTGACAGGAATGATTGAA 2412  
Qy ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840  
Db 2413 ACCGACGAATGACTGATTTGCCAACAAAGATAGCAAGAACTTTAGCATGCTGGAAG 2472  
Qy IleAlaThrGluAlaLeuGluAsnIleAtrGThrIleValSerLeuThrArgGluLysAla 860  
Db 2473 ATAGCAACTGAAGCTTTGGAGATATACGTACTATAGTGTCAATTAACAGGGAAGGCC 2532  
Qy PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880  
Db 2533 TTCGAGCAAAATGTTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2592  
Qy GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900  
Db 2593 CAGATTATTGGAAGCTGTTATGATTCAGCCATGCTCTTATATATTGTCCTATGACGCA 2652  
Qy GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920  
Db 2653 GGGTTTCGATTGGAGCTATTATTAATCAAGCTGGAGATGACCCAGAGGCGCATGTTTC 2712  
Qy IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940  
Db 2713 ATAGTTTCTACTGCAATTCGATATGAGCTATGCGCATGCGCAAAACGCTCGTTTGGCT 2772  
Qy ProGluTyrSerLysAlaLysSerGlyAlaAlaHisIleuPheAlaLeuGluLysLys 960  
Db 2773 CCTGAATATTCGAAGCCAAATCGGGGGCTGCGCATCTGTTGCTTTGTTGGAAGAAAGAA 2832  
Qy ProAsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeu 980  
Db 2833 CCAATATAGACAGCCGACATCAAGAGGGGAAAGCCACACATGTGAAGGGGATTTA 2892  
Qy GluPheArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGly 1000  
Db 2893 GAGTTTCGAGAGTCTCTTCTCTATCCATGTCGCCAGATGTTTTCATCTCTCCGTGGC 2952  
Qy LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020  
Db 2953 TTAATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGAGACGCGCTGTGG 3012  
Qy LysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040  
Db 3013 AAGAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGAGCAAGTGTG 3072  
Qy PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060  
Db 3073 TTTGATGGTGTGATGCAAAAGAAATGAATGTACAGTGGCTCCGTTCCCAAAATAGCAATC 3132  
Qy ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080  
Db 3133 GTTCCCTCAAGAGCGCTGTGCTCTTCACTGAGCATTCCTCAGAACATTCGCTATGGTGAC 3192  
Qy AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100  
Db 3193 AACAGCGTGTGGTGCCTATTAGATGAGATCAAGAGCCGCAATGCGCAATATATCCAT 3252  
Qy SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120  
Db 3253 TCTTTTATTGAAGTCTCCCTGAGAAATCAACACACAACTTGAGCTGAAAGGAGCACAG 3312  
Qy LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 1140  
Db 3313 CTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTTCCAAAAACCCAAA 3372

Qy 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAspSerGluLysValValGln 1160  
Db 3373 ATTTTATTGTTGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAGGTGGTTTCAG 3432  
Qy 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180  
Db 3433 CATGCCCTTTGATAAAGCCAGGACGGGAAGACATGCTTAGTGTCTACTCACAGGCTCTCT 3492  
Qy 1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200  
Db 3493 GCATTCAGAACCGCATTTGATAGTGGTTCTGCCAATGGAAGATAAAGGAACAAGGA 3552  
Qy 1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 3553 ACTCATCAAGAGCTCTCGAATAATCGAGAAATCGAGACATATATTTTAAAGTTAGTGAATGCACAGTCA 3612  
Qy 1221 ValGln 1222  
Db 3613 GTGCAG 3618  
RESULT 3  
US-09-873-409-12  
; Sequence 12, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 3177  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (198)..(198)  
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)  
US-09-873-409-12  
Alignment Scores:  
Pred. No.: 0 Length: 3177  
Score: 5328.00 Matches: 1058  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 86.51% Indels: 0  
DB: 10 Gaps: 0  
US-09-873-409-5 (1-1222) x US-09-873-409-12 (1-3177)  
Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184  
Db 1 ATGGTCACTCTCATTCACCAAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60  
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204  
Db 61 GAAGAAAGTCTTGTCATCAATCCGAACAGTCATAGCCCTTTAGGGCCCGGAGGAAGAACTT 120  
Qy 205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224  
Db 121 CAAAGGCTCTTCTCTTAAATATAACAGATATGCTTGTTTATTTTCCCCAGTGGCTA 180  
Qy 225 LeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244  
Db 181 CTAAGTTGTTCTGTTNTTGTAAAGTATACACAGAACTCTCAAGATGCAAGAGGATTTT 240  
Qy 245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264

Db 241 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTACTCTTTTATGAAT 300  
Qy 265 GlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAenGlyGluProGly 284  
Db 301 GGAACTTATGGACTTGTCTTTTGGTATGGAACCTCTCTTGATTTCTTAATGAGAACCTTGA 360  
Qy 285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304  
Db 361 TATACCATCGGACTGTCTCTGCTGTTTTCTTTAGTGTATTCATACAGCACTATTGCATT 420  
Qy 305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324  
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCTGCTTTCATATT 480  
Qy 325 PheGlnValIleLeuPheLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344  
Db 481 TTCAGGTTTATTGATAAGAACCCAGATATAGATAACTTTTCCAGCTGGATGATATAACCT 540  
Qy 345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364  
Db 541 GAATCCATAGAGGAACGTGGAAATTTAAAAATGTTCTTTCAATTATCCATCAAGACCA 600  
Qy 365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384  
Db 601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCCTG 660  
Qy 385 ValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404  
Db 661 GTCGGTCTCAATGGCAGTGGGAAGATGACGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720  
Qy 405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424  
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGGGCAT 780  
Qy 425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 444  
Db 781 TATCGAGACCATATTGGAGTGGTGTAGTCAAGAGCTGTTTTGTTCGGGACCAACCATCAGT 840  
Qy 445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464  
Db 841 AACAAATATCAAGTATGGACGAGATGATGTACTGATGAAGAGATGGAGAGCAGCAAGG 900  
Qy 465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484  
Db 901 GAACCAATCGGTATGATTTATCATGGAGTTCTCTAATAAATTTAATACATGGTAGGG 960  
Qy 485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504  
Db 961 GAAAAAGAGAGCTCAAAATGAGTGGAGGGGCAGAAACAGAGGATCGCAATTTGCTGCGCTTA 1020  
Qy 505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524  
Db 1021 GTTCGAACCCCAAGATCTGATTTTATGATGAGGCTACGCTGCGCCCTGGATTCAGAAAGC 1080  
Qy 525 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 544  
Db 1081 AAGTCAGCTGTTCAGCTGCTGAGAGCGCAGCAAGGTCCGACTACAAATCGTGGA 1140  
Qy 545 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 564  
Db 1141 GCACACCGACTTTCTACTATTTCGAATGCGAGATTGATTTGTGACCCCTAAAGGATGGAATG 1200  
Qy 565 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 584  
Db 1201 CTGGCGGAGAAAGGAGCAGCATGCTGGAATTAATGGCAAAACAGGCTCTATATTATTCACCT 1260  
Qy 585 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 604  
Db 1261 GTGATGTTCAGAGATATTAATAAGCTGATGACAGATGGAGTCAATGACATATTCTACT 1320  
Qy 605 GluArgLysThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624  
Db 1321 GAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy 625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644  
Db 1381 GACAAGGCTGAGNAATCCACCAATCTAAAGAGATAAGTCTTCTCGAAGTCTCTCTATTATTA 1440  
Qy 645 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 664  
Db 1441 AAAATTTTAAAGTTAAACAAGCTGAATGCCTTTTGTGGTCTCTGGGACATTTGGCTTCT 1500  
Qy 665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684  
Db 1501 GTTCTAAATGGAACTGTTTCATCCAGTATTTTCATCATCTTTTGCAAAAATTAACCATG 1560  
Qy 685 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704  
Db 1561 TTTGGAAATAATGATATAAACACCATTAAGACATGATGCAGAAATTTATTCCATGATATTC 1620  
Qy 705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724  
Db 1621 GTCAATTTTGGGTGTTATTTGCTTTGTCACTTATTTTCATGCAGGGATTTATTTACGGCAGA 1680  
Qy 725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744  
Db 1681 GCAGGGGAAATTTTAACGATGAGATTAAAGACACTTGGCCCTTCAAGGCCATGTTATATCAG 1740  
Qy 745 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764  
Db 1741 GATATTGCTGTTTGTATGAAAGGAAACACACAGCAGGAGCTTGACACAAATATTAGCC 1800  
Qy 765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784  
Db 1801 ATAGATATAGCAAAATTCAGGAGCAACAGGTTCCAGGATTTGGCGTCTTAACACAAAAT 1860  
Qy 785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804  
Db 1861 GCAACTTAACATGGAGCTTTTCAGTTATCTTCTTATATATGATGGAGTGGAGATGACATTC 1920  
Qy 805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824  
Db 1921 CTGATTTCTGAGTATTTGCTTCAGTACTTGGCGTGACAGGAATGATGAAACCCGAGCAATG 1980  
Qy 825 ThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGlu 844  
Db 1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGAGATAGCAACTGAA 2040  
Qy 845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864  
Db 2041 GCTTTGGAGAAATATAGTACTATAGTGTCTATTAAACAAGGAGAAAGCCCTTCGAGCAATG 2100  
Qy 865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884  
Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGACACAGATATTGGA 2160  
Qy 885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904  
Db 2161 AGCTGTTATGCATTCAGCCATGCTTTATATATTTTGCCTATGCTATGCAGCAGGGTTTCGATTT 2220  
Qy 905 GlyAlaTyrIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924  
Db 2221 GGAGCCTATTAAATCAAGCTGGACGAATGACCCAGAGGGCATGTTTCATAGTTTACT 2280  
Qy 925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944  
Db 2281 GCATTTGCATATGGAGCTATGGCCATCGGAAACCCCTCGTTTGGCTCTCGAATATTC 2340  
Qy 945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964  
Db 2341 AAAGCCAAATCGGGGCTGCGCATCTGTTTGCCTTGTGGAAAAAGAAACCAATATAGAC 2400  
Qy 965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984  
Db 2401 AGCCCGAGTCAAGAGGGAAGAAAGCCAGACACATGTGAAGGGGAATTTAGAGTTTCAGAA 2460

QY	985	ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer	1004
Db	2461	GTCTCTTTCTTCATCCATGTCGCCAGATGTTTTCATCTCCGTCGGCTTATCCCTCAGT	2520
QY	1005	IleGluArgGlyLeuThrValAlaPheValGlySerSerGlyCysGlyValysSerThrSer	1024
Db	2521	ATTGACGCGAGGAAGACAGTAGCATTTGTGGGGAGCAGCGCTGTGGGAAAGACATCTCT	2580
QY	1025	ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal	1044
Db	2581	GTTCAACTTCTGCAGAGACTTTATGACCCCGTCGCAAGGACAAGTGCTGTTTGATGGTGTG	2640
QY	1045	AspAlaLysGlnLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu	1064
Db	2641	GATGCAAAAAGAAATTGAATGTACAGTGGCTCGGTTCCCAATAGCAATCGTTCCTCAAGAG	2700
QY	1065	ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal	1084
Db	2701	CCGTGCTCTTCAACTGACGACATGTCTGAGAACATCGCCTATGGTGACAAACGCCGTGTG	2760
QY	1085	ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu	1104
Db	2761	GTGCCATTAGATGAGATCAAGAGCGCGCAATGCAGCAATATCCATCTTTTATTGAA	2820
QY	1105	GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly	1124
Db	2821	GGTCTCCCTGAGAAATACAAACACACAAGTTGGACTGAAAGGACACACAGCTTTCTGGCGGC	2880
QY	1125	GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu	1144
Db	2881	CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAAATTTTATTGTTG	2940
QY	1145	AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp	1164
Db	2941	GATGAGCCCACTTCAGCCCTCGATATGACAGTGAGAGGTGGTTGACGATGCCCTTGAT	3000
QY	1165	LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn	1184
Db	3001	AAAGCCAGGACGGGAAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC	3060
QY	1185	AlaAspIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGlu	1204
Db	3061	GCAAGATTGATAGTGGTTCTGCACAAATGGAAGATAAAGGAACAAGGAACCTCATCAAGAG	3120
QY	1205	LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln	1222
Db	3121	CTCCTGAGAAATCGAGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG	3174
RESULT 4			
US-09-873-409-10			
; Sequence 10, Application US/09873409			
; Patent No. US20020037522A1			
; GENERAL INFORMATION:			
; APPLICANT: Frank, Markus			
; APPLICANT: Sayegh, Mohamed			
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein			
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof			
; FILE REFERENCE: 81994/268611			
; CURRENT APPLICATION NUMBER: US/09/873,409			
; CURRENT FILING DATE: 2001-06-05			
; NUMBER OF SEQ ID NOS: 19			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 10			
; LENGTH: 2856			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-873-409-10			
Alignment Scores:			
Pred. No.:	0	Length:	2856
Score:	4561.50	Matches:	912
Percent Similarity:	97.96%	Conservative:	0
Best Local Similarity:	97.96%	Mismatches:	1

Query Match:	74.06%	Indels:	18
DB:	10	Gaps:	1
US-09-873-409-5 (1-1222) x US-09-873-409-10 (1-2856)			
QY	293	ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheG1	312
Db	114	GTMTTCTTTTAGTGAATCCATAGCAGTTATTGTCATTGGGAGCAGCAGTCCCTCATTTATT	171
QY	312	uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr	332
Db	172	-----GATAAGAAACC	182
QY	332	oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValG1	352
Db	183	CAGTATAGATAAATCTTCCACAGCTGGATATAAACCTGAATCCATAGAAGAACTGTGA	242
QY	352	uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe	372
Db	243	ATTTAAAAATGTTCTTTCAATTTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT	302
QY	372	uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyL	392
Db	303	GAACTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTCGGTCTCAATGGCAGTGGGA	362
QY	392	sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa	412
Db	363	GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGATGATGGCTTTATCATGTGT	422
QY	412	lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa	432
Db	423	GGATGAGAAATGACATCAGAGCTTTTAAATGTGGCGCATTTATCGAGACCATATTTGGAGTGT	482
QY	432	lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAa	452
Db	483	TAGTCAAGAGCCTGTTTTGTTCCGGACCACTCATGATTAACAATATCAGTATGGACGAGA	542
QY	452	pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheI	472
Db	543	TGATGTGACTGATGAAGACATGGAGAGACAGCAAGGGGAAGCAAAATGCGTATGATTTTAT	602
QY	472	eMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerG1	492
Db	603	CATGGAGTTTCTTAATAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAAATAGTGG	662
QY	492	YGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu1	512
Db	663	AGGCGAGAAACAGAGATCGCAATTGCTCGTCCCTTAGTTTGGAAACCCCAAGATTCTGAT	722
QY	512	eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe	532
Db	723	TTTAGATGAGGCTACGTCGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGTCGACT	782
QY	532	uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr	552
Db	783	CGAGAAAGCGCAGCAAAAGTTCGGACTACAATCGTGGTAGCACACCCGACTTTCTACTATT	842
QY	552	gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl	572
Db	843	AAGTGCAGATTGATTGTGACCCCTAAAGGATGAATGCTGGCGGAGAAAGGAGACATGC	902
QY	572	aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy	592
Db	903	TGNACTAATGGCAAAACAGGCTCTATATTATTCACCTTGATGTCACAGATATTAAAA	962
QY	592	sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr	612
Db	963	AGCTGATGAACAGATGAGTCAATGACATATTTCTACTGAAAGAAAGACCAACTCAGCTTC	1022
QY	612	oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSerThrG1	632
Db	1023	TCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCATTTCAGAGGCTGAGGAATCCACCCA	1082

QY 632 nSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 652  
DB 1083 ATCTAAAGAGATAGTCTTCTGAAGTCTCTCTATTAAAAATTTAAAGTTAAACAGCC 1142  
QY 652 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 672  
DB 1143 TGAATGGCCCTTTTGTTCTGGGGACATTTGGCTCTCTTTCTAAATGGAACTGTTCAATC 1202  
QY 672 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 692  
DB 1203 AGTATTTTCATCATCTTTGCATAAATTAATTAACCATGTTTGGAAATTAATGATAAACCCAC 1262  
QY 692 rLeuLysHisAspAlaGluLysIleSerMetIlePheValIleLeuGlyValIleCysPh 712  
DB 1263 ATTAAGCATGATGCAGAAATTTATTCATGATATTCGTCAATTTGGGTGTTATTTGCTT 1322  
QY 712 eValSerTrpPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 732  
DB 1323 TGTCAGTTATTTTCATGAGGATTTATTTACGGCAGACAGGGGAAATTTTAAACGATGAG 1382  
QY 732 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 752  
DB 1383 ATTAAGACACTTGGCCCTCAAGCCATGTTATATCAGATATTCGCTGGTTTGATGAAA 1442  
QY 752 sGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnI 772  
DB 1443 GGAACACAGCACAGGAGGCTTGACAAATATTTAGCCATAGATATAGCAAAATTTCAAGG 1502  
QY 772 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 792  
DB 1503 AGCAACAGGTTCCAGGATTTGGCTCTTAACACAAAATGCAACTAACATGGGACTTTTCAGT 1562  
QY 792 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 812  
DB 1563 TATCATTTCTTTATATGATGGAGATGACACTTCTGATTCGAGTATTCCTCAGT 1622  
QY 812 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 832  
DB 1623 ACTTGGCGTGACAGGAATGATTGAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682  
QY 832 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrII 852  
DB 1683 GCAAGAACTTTAAGCATCTGCGAAAGATAGCAACTGAAAGCTTTGGAGAATATACGTACTAT 1742  
QY 852 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 872  
DB 1743 AGTGCTATTAAACAGGGAAGAAAGCCCTTCGAGCAATGATGTAAGAGATGCTTCAGACTCA 1802  
QY 872 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 892  
DB 1803 ACACAGAAATACCTCGAAGAAAGCACAGATTTATTGGAAGCTGTTATGCATTCAGCCATGC 1862  
QY 892 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 912  
DB 1863 CTTTATATATTTTCCCTATGATGAGGAGGTTTCGATTTGGAGCCCTATTAAATCAAGCTGG 1922  
QY 912 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 932  
DB 1923 ACGAATGACCCAGAGGCGATGTTCTAGATTTTACGTCAATTCATATGAGCTATGGC 1982  
QY 932 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 952  
DB 1983 CATCGGAAACCGCTCTTTTGGCTCTCTGATATTTCCAAAGCCAAATCGGGGCTGGCA 2042  
QY 952 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLy 972  
DB 2043 TCTGTTTCCCTTGTGGAAAGAAACCAAAATATATACAGCGCGAGTCAAGAGGGGAAAAA 2102  
QY 972 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysAr 992  
DB 2103 GCCAGACATGTGAAGGAAATTTAGAGTTTCGAGAGATCTCTTCTTCATTCATGCTGC 2162  
QY 992 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 1012

DB 2163 CCAGATGTTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGGAAAGACAGTAGC 2222  
QY 1012 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1032  
DB 2223 ATTTGTTGGGAGCAGCGGCTGTGGGAAAGACATCTCTGTTCACCTTCTCGAGAGACTTTA 2282  
QY 1032 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1052  
DB 2283 TGACCCCGTGCAAGGACAGTGTCTGTTCATGGTGTGGATGCAAAAGAAATTTGAATGTACA 2342  
QY 1052 nTrpLeuArgSerSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1072  
DB 2343 GTGGCTCCGTTCCTCAATAGCAATCGTCTCAAGAGCCCTGTCTCTTCAACTGCGAGCAT 2402  
QY 1072 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1092  
DB 2403 TGCTGAGAACATCGCTATGTGTGACAAACCGCTGTGGTGCCATTAGATGAGATCAAGA 2462  
QY 1092 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1112  
DB 2463 AGCCGCAATGCAGCAATATCCATTTCTTTTATTGAAAGTCTCCCTGAGAAATACAACAC 2522  
QY 1112 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1132  
DB 2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGCCAGAAACAAAGACTAGCTATTGC 2582  
QY 1132 aArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAs 1152  
DB 2583 AGGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642  
QY 1152 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1172  
DB 2643 TAATGACATGAGNAGTGTGTCAGCATGCCCTTGATTAAGCCAGGACGGGAGGACATG 2702  
QY 1172 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1192  
DB 2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACCGCAGATTTTGATAGTGTCTGCA 2762  
QY 1192 sArgGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1212  
DB 2763 CAATGGAAGATGAAAGAAACAAAGAACTCATCAAGAGCTCTCTGAGAAATCGAGACATATA 2822  
QY 1212 rPheLysLeuValAsnAlaGlnSerValGln 1222  
DB 2823 TTTTAAGTTAGTGAATGCACAGTCAGTCAGCAG 2853

## RESULT 5

US-09-866-866A-1  
; Sequence 1, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 3860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-866-866A-1

Alignment Scores:

Pred. No.: 4,44e-314 Length: 3860  
Score: 3530.50 Matches: 682  
Percent Similarity: 74.00% Conservative: 246  
Best Local Similarity: 54.39% Mismatches: 263  
Query Match: 57.32% Indels: 63  
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-866-866A-1 (1-3860)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 151 ATGGTGGTGGGAACCTTGGCTGCATCATCCATGGGCTGGACCTTCTCTCATGATGCTG 210  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAAAGATCTGATGTCA 270  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 271 AACATCACTAATAGAAATGATATCAATGATACAGAGGTTCTTCAATGATCTGGAGGAAGAC 330  
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 331 ATGACACAGATATCCCTATTATTACAGTGAATTTGGTCTGGGGTGTGGTGTGCTGTAC 390  
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 391 ATTCAAGTTTCATTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAATTTAGAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGCTATAATCGCAGGAGATAGGCTGGTGTGATGTCACAGCTGTGGGGAG 510  
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 511 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAAGGATATTGGTGACAAAT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGCTTTTCAGTCAGTGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAGCT 630  
Qy 141 GlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GGTGTGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGACATGTGTCAGCTGT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 691 GTCGTGGCAAGATATCTTCTTACTGTATAAAGAACTCTTAGCGTATGCAAAAGCT 750  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTTAGAACTGTGATTGCAATTTGGAGGACAA 810  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 811 AAGAAAGAACTTGAA----- 825  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 826 -----AGGTACAAACAAAATTTAGAAAGAA 849  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 850 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAAATTTCTATAGTGTGCTCTTC 909  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
Db 910 CTGCTGATCTATGATCTTATGCTCTGGCTCTCTGGTATGGGACCACTTTGGTCTCTCA 969  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 970 GGGGAA-----TATTCTATTGGACAAAGTACTCACTGTATTCTTTCTGTGATTAAATGGG 1023

Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1024 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTGCAAAATCGAAGAGAGCA 1083  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1084 GCTTATGAATCTTCAAGATAATTGATAAAGCCAGTATTTCACAGCTATTTCGAAGAGT 1143  
Qy 341 GlyTyrLysProGluSerIleGluThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1144 GGGCACAACACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTTCACTTTCAGTTAC 1203  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1204 CCATCTCGAAAAGAAATTAAGATCTTGAAGGCGCTGAACCTGAAGGTGAGAGTGGGAC 1263  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGTATGTCAG 1323  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1324 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTATGGACAGGATATTAGGACCATA 1383  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1384 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGCC 1443  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1444 ACCAGATAGCTGAAACCAATTCGTATGGCCGTGAAATGTCCACATGATGATGATGAG 1503  
Qy 461 ArgAlaAlaArgGluAlaAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1504 AAAGCTGTCAAGAAAGCCAAATGCCATGATGATTTATCATGAACTGCCTATAAATTGAC 1563  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1564 ACCCTGGTTGGAGAGAGAGGGGCCCATGTTGAGTGGTGGCAGAAAGCAGAGGATCGCAT 1623  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCCTCTGCTGGATGAGGCCACGTCAGCTTG 1683  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1684 GACACAGAAGCGAAGCAGTGGTTCAGTGGTCTCTGGATAGGCCAGAAAGGTCCGACC 1743  
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1744 ACCATTGTGATAGTCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGTTTC 1803  
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1804 GATGATGGAGTCAATTGTGGAGAAAGGAAATCATGATGAATCATGAAAGAAAGGCATT 1863  
Qy 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
Db 1864 TACTTCAAACTTGTCAATGACAGCAGCAGGAAATGAAGTTGAATTAGAAATGCGACT 1923  
Qy 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 1924 GATGAATCCAAAAGTGAATTTGATGCTTTGGAAATGCTCTTCAATGATTCAAGATCCAGT 1983  
Qy 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
Db 1984 CTAATAAGAAAAGATCAACTCGTAGGAGTGTCCGTGGATCAACAGCCCAAGACAGAAAG 2043  
Qy 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647  
Db 2044 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTCTTTGGAGGATATG 2103  
Qy 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667



```
Db 2104 AAGCTAAATTAAGTGAATGCCTTATTTGTTGTTGCTGATTTTGTGCGCATATAAAT 2163
Qy 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db 2164 GGAGGCTGCACACGAGCATTTGCATAATATTTTCAAGAAATATAGGGGTTTTCACAGA 2223
Qy 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db 2224 ATTGATGATCCTGAAACAAAACGACAGAAATAGTAACCTTTTTCACATATTGTTTCAGCC 2283
Qy 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db 2284 CTTGGAATATTCTTTTATTACATTTTCTTCAAGGTTTTCATTTGCGCAAGAGCTGGA 2343
Qy 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db 2344 GAGATCCTCACCAAGCGGCTCCGATACATGTTTTCGATCCATGCTCAGACAGGATGTG 2403
Qy 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db 2404 AGTTGGTTTGATGACCTTAAACACCACTGGAGCATTTGACTACCGGCTCGCCAAATGAT 2463
Qy 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTCTGCTTAATTCACCAAGATATAGCA 2523
Qy 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806
Db 2524 AATCTTGGGACAGGAATAATATATCTCTCATCTATGTTGGCACTTAACACTGTTTACTC 2583
Qy 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 2584 TTAGCAATTGTACCATCAITTCGAATAGCAGGAGTTGTTGAAATGAAATGTTGCTCGGA 2643
Qy 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 2644 CAAGCACTGAAGATAAGAAAGAACTAGAAGGCTCTGGGAAGATCGCTACTGAAGCAATA 2703
Qy 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 2704 GAAACATTCGGAACCGTTGTTCTTTCACTCAGGAGCAGGAGTTTGAACATATATATGCT 2763
Qy 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 2764 CAGAGTTTGAGGTTGACCATACAGAACTCTTGAGGAAGACACACATCTTTGGAATTACA 2823
Qy 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 2824 TTTTCCCTTACCACCGCAATGATGATATTTTCTATGCTGATGTTTCCCGTTTGGAGCC 2883
Qy 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 2884 TACTTGTGGCACATAAACAATCATGAGTTTGTGAGGATGTTCTGTAGTATTTTCAGCTGTT 2943
Qy 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 2944 GTCTTTGGTGCATGGCGGGGCAAGTCAGTTCAITTTGCTCTGACTATGCCAAGCC 3003
Qy 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3004 AAAATATCAGCAGCCACATCATCATGATCATTAATAAACCCTTTGATTCAGACGCTAC 3063
Qy 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3064 AGCAGGAAGGCTTAATGCCGGAACACATCTGGAAGAAATGTCAATTTGTGTAAGTTGTA 3123
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3124 TTCAACTATCCACCCGACCGGACATCCAGTGCTTCAGGAGCTGAGCCTGGAGGTGAAG 3183
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
Db 3184 AAGGGCCAGCAGCTGCTCTGTTGGCCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3243
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAla 1046
Db 3244 CTCCTGGAGGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGCTGATGGCAAGAAATA 3303
Qy 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3304 AAGCGACTGAATGTTTCAGTGGCTCCGAGCACCTGGGCATCGTGTCCAGGAGCCATC 3363
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3364 CTGTTTTCAGTGCAGCATTGTGAGAACTTGGCTATGGAGACAACAGCCGGGTGTCA 3423
Qy 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3424 CAGGAAGATCTGTAGGCGAGCAAGAGGCGCAACATATACATGCTTCATCGAGTCACTG 3483
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3484 CCTAATAATATAGCACTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGTGTGCCAGAAA 3543
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3544 CAACGATTCGCCATAGCTCGTGCCTTGTAGACAGCCTCATATTTGCTTTTGGATGAA 3603
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3604 GCCACGTCAGCTCTGATACAGAAAGTGAAGGTTGTCCAGAAGCCCTGGACAAAGCC 3663
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3664 AGAGAAGCGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3723
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGluLeu 1206
Db 3724 TTAATAGTGTGTTTCAGAAATGGCAGATGCAAGGAGCATGGCAGCATCAGCAGCTGCTG 3783
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 6
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores: 1.51e-313 Length: 4317
Pred. No.: 3525.50 Matches: 680
Score:
```



Percent Similarity: 73.05% Conservative: 239  
Best Local Similarity: 54.05% Mismatches: 270  
Query Match: 57.24% Indels: 69  
DB: 9 Gaps: 9

US-09-873-409-5 (1-1222) x US-10-044-671-1 (1-4317)

Qy 1 MetIleLeuGlyIleLeuAlaSerIleuValAenGlyAlaCysLeuProLeuMetProLeu 20  
Db 217 ATGTTGGTGGGACAAATGCTGCCATCATCCATGAGCTGCACCTCCTCATGATGCTG 276  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 277 GTTTTGGAAACATGACATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 336  
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 337 CCAGTTATATTAATGAAGATATTACGAACATATACACACATTTTCATCAACCATCGGAG 396  
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
Db 397 GAGGAAATGACCACGATGATCCCTATTATTACAGTGGGATCGGTGCTGGTGGCT 456  
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
Db 457 GCTTATACATCCAGGTTTCATCTGCTGCTGGCAGCAGGAAGACAGATCTCAAAATTAGA 516  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99  
Db 517 AACCAATTTTTTCATGCTATATCGACGAGGAGATTGGCTGGTTTACGTGATGACGTT 576  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 577 GGGGAGCTTAACACCGGCTCACAGACGATGCTCCAAATCAATGAAGGAATGCGCAG 636  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 637 AAGTTTGAATGTTCTTCAATCAATAGCAACATTTTCCACCGTTTATAGTGGGGTTT 696  
Qy 139 ValLysGlyTrpLysLeuLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 697 ACACCTGGTTGGAAGTAACTACCTTTGATTTTGGCCATCAGCCCTGTTCTGGACCTTCA 756  
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 757 GCGCCATCTGGGCAAGATACCTATCTTCTTACTGATAGAAGAACTCTTGGCCTATGCA 816  
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 817 AAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 876  
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218  
Db 877 GGACAAAGAAAGAACTTGAA----- 897  
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeu 238  
Db 898 -----AGTACAAACAAATTTA 915  
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
Db 916 GAAGAAGCTAAAGAATTTGGATAAAGAAAGCTATCAGGCCCAACATTTCTATTGGTGC 975  
Qy 259 valTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278  
Db 976 GCTTTCTTATGATCATGATCATATGCTCTGGCTTTCTGGTATGGGACCTCTTTGGTC 1035  
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
Db 1036 CTCTCCAGTGAA-----TATACTATGGACAGGTACTCACTGCTCTCTTTCTGATTA 1089  
Qy 299 HisSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
Db 1089 -----TATACTATGGACAGGTACTCACTGCTCTCTTTCTGATTA 1089

Db 1090 ATTGGGGCTTTTAGTATTGGACAGCATCCCAACGANTTGAAGCATTTGCAAAACCAAGA 1149  
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
Db 1150 GGCAGAGCTTATGAATCTTCAAGATAATTGAACAATAAACCAAGCATTCACAGCTATTGC 1209  
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358  
Db 1210 AAGAGTGACATAAACAGATAATATTAAAGGAATTTTGAATTTCAAAATGTTCACTTC 1269  
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
Db 1270 AGTTACCTCTTCGAAAAGAAGATTAAAGATCTTAAGGGTCTCAACTGAAGGTTCAAGT 1329  
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
Db 1330 GGCAGACAGTGGCGCTGTTGGGAACAGTGGCTCGGGAAGAGCAGCACCGCTGCAGCTG 1389  
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
Db 1390 ATGCAGAGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGGACAGACATTAGG 1449  
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
Db 1450 ACCATAAATGTAAAGCATCTTCGGGAATTTACTGGTGTGGTGGTGGTGGTGGTGGT 1509  
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458  
Db 1510 TTTGCCACCATGATAGCTCAAAACATTCGCTATGCGCGGAAATGTCAACATGGATGAG 1569  
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProLeuLys 478  
Db 1570 ATTGAGAAAGCTGTTAAAGGAAGCAATGCTATGATTTTATCATGAACATGACCTAATAA 1629  
Qy 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
Db 1630 TTTGACACTCTGGTGGAGAGAGGGGGCGCGCTGAGTGGTGGGACAGAAACAGAAATC 1689  
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
Db 1690 GCATTTGCTCGGGCCCTGGTTCGCAACCCCAAGATTTCTTCTGATGAGGCAACGTCA 1749  
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGly 538  
Db 1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGGATAAGGCCAGAAAGGC 1809  
Qy 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558  
Db 1810 CGGACTACCATTTGTATAGCTCATCGTTTGTCTACAGTTTCGTAATGCCGATGTCATGCT 1869  
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578  
Db 1870 GGTTTTGTATGATGGAGTCAATTTGGAGAAAGGAATCATGATGAACATCATGAAGAGA 1929  
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598  
Db 1930 GGCATTTACTTCAAACTTGTCAATGCGAG---ACAAGAGGAAATGAAATTTGAGTTAGAA 1986  
Qy 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
Db 1987 AATGCCACTGGTGAATTCACAAAAGTGAAGTGAAGTGGCTTGGAAATG---TCTCCAAAGAT 2043  
Qy 619 IleLysSerAspPheIle----- 624  
Db 2044 TCAGGGTCCAGTTTAAATAAAGAGATCACTCCGAGGAGTATACATGCACCACCAAGGC 2103  
Qy 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643  
Db 2104 CAAGACAGAAAGCTTGGTACAAAAGAGAGGACTTGAATGAGAATGTTCTTTCAGTTTCTC 2163  
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663  
Db 2164 TGGAGGATTTCTGAAGCTGAACCTCACTGAATGCGCTTATTTTGTGGTGTGTTATTTGT 2223

QY 664 SerValLeuAenGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683  
Db 2224 GCTATTATAACGGAGCGCTGCAACACAGCATTTTCAATATATTTCAGGATATTAGGG 2283  
QY 684 MetPhe---GlyAenAenAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702  
Db 2284 ATCTTTACCGAGATGAGGATCTCTGAAACAAACGACAGAAATAGTAACTATGTTCTGTA 2343  
QY 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722  
Db 2344 TTGTTTCTAGTCTTGAATATTCTTTATTACATTTTCTCCAGGCTTCAATTT 2403  
QY 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisIleLeuAlaPheLysAlaMetLeu 742  
Db 2404 GGCAGAGCTGGGAGATCTCCTAAGCGCTCGATACATGTTTTCAGATCCATGCTG 2463  
QY 743 TyrGlnAspIleAlaTyrPheAspGluLysGluAenSerThrGlyGlyLeuThrThrIle 762  
Db 2464 AGACAGATGTCAGCTGTTGATGACCTTAAACACCAACCACTGGAGCATTTGACAAACGAG 2523  
QY 763 LeuAlaIleAspIleAlaGlnIleGlnGlyValaThrGlySerArgIleGlyValLeuThr 782  
Db 2524 CTGCGCAATGATCGGCTCAAGTTAAAGGGCTATAGTTCCAGGCTTGCTGCAATTACC 2583  
QY 783 GlnAsnAlaThrAenMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 802  
Db 2584 CAGNATATAGCAATCTGGGACAGGCAATATTATATCTTATCTATGTTGGCAATTA 2643  
QY 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822  
Db 2644 ACATTTTACTCTTAGCAATTTGATACCATCATTCGATAGCAGGAGTTGTTGAAATGAAA 2703  
QY 823 AlaMetThrGlyPheAlaAenLysAspLysGlnLeuLysHisAlaGlyLysIleAla 842  
Db 2704 ATGTTGTCTGGACAGCACTGAAAGATAAGAAAGAGCTAGAAAGAGCTGGGAGATTGCT 2763  
QY 843 ThrGluAlaLeuGluAenIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862  
Db 2764 ACAGAGCCATCGAATCTCCGAACTGTTGTTCTTTGACTCGGAGAGCAAGTTGAA 2823  
QY 863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAenThrSerLysLysAlaGlnIle 882  
Db 2824 TACATGATGACAGAGTTTGCAAGTACCACAGCACTTTTGAGGAAAGACACATC 2883  
QY 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902  
Db 2884 TTCGGGCTCTCATTTCTATCCAGGCAATGATGATTTTCTATGCTGGCTGTTTC 2943  
QY 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922  
Db 2944 CGGTTTGGTGCCTACTTGGTGCAATGAGTTTCACTTTCAGGATGTTCTTTTGTA 3003  
QY 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942  
Db 3004 TTCTCAGTATTGCTTTGTTGGTCCATGCGAGTGGGCGGAGTCACTTATTTGCTCTGAC 3063  
QY 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962  
Db 3064 TATGCCAAGCCAAAGTATCAGCAGGCCACGCTCATGATCATTTGAAAGAGCCCTCTG 3123  
QY 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyValAenLeuGluPhe 982  
Db 3124 ATTGACAGCTACAGCCCTCAGCGGCTCAGCCAAATAGTTGGAGGAAATGTCACATTT 3183  
QY 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002  
Db 3184 AATGAGGTCGGTTCAACTATCCCACTCGACAGACATCCCGGTGCTCCAGGGCTGAGC 3243  
QY 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022  
Db 3244 CTCGAGGTGAAGAGGGCCAGACGCTGCGCCCTCTGAGTAGCAGTGGCTGTGGGAAGAGC 3303

QY 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042  
Db 3304 ACAGTTTGTTCAGCTCTCAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTAAATTGAT 3363  
QY 1043 GlyValAspAlaLysGluLeuAenValGlnTyrLeuArgSerGlnIleAlaIleValPro 1062  
Db 3364 GGCAGAGATGAAGACCTGAATGTCAGTGGCTCCGAGCACACCTGGGCACTGTTGCT 3423  
QY 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAenIleAlaTyrGlyAspAsnSer 1082  
Db 3424 CAGGAGCCCATCTGTTTTCAGTGCAGCATTTGCCGAGAACATTCCTATGAGAGCAACAGC 3483  
QY 1083 ArgValValProLeuAspGluIleLysGluAlaAaAsnAlaAlaAsnIleHisSerPhe 1102  
Db 3484 CGGTCGTATCATCATGAGAGATTATTCAGGCGAGCCCAAGGAGCCCACTACACCTTC 3543  
QY 1103 IleGluGlyLeuProGluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122  
Db 3544 ATCGAGACATCTCCTGAGAAATACAAACAGAGTAGGAGACAAAGAACCCAGCTCTCT 3603  
QY 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142  
Db 3604 GGTGGCAGAAACAGCGCATTCGTCATAGCTCGGCTCTGTTAGACAGCCTCATATTTC 3663  
QY 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162  
Db 3664 CTTTTCGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAGAGGCC 3723  
QY 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182  
Db 3724 CTGCAAAAGCCAGAGAGCGCCACCTGCTGATGTCGCCACCGCTTGTCCACCATC 3783  
QY 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202  
Db 3784 CAGATGTCAGATTATATAGTGGTTCAGATGGCAAGTCAAGGAGCATGCGACACAT 3843  
QY 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 3844 CAACAGCTGCTGGCTCAGAAAGGCATCTATTTTCCATGATCATGTCAGTGTCCAGGCT 3897  
RESULT 7  
US-09-866-866A-3  
; Sequence 3, Application US/09866866A  
; Patent No. US2002010244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 3860  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-09-866-866A-3  
Alignment Scores:  
Pred. No.: 2,98e-313 Length: 3860  
Score: 3521.50 Matches: 681  
Percent Similarity: 73.92% Conservative: 246  
Best Local Similarity: 54.31% Mismatches: 264  
Query Match: 57.18% Indels: 63  
DB: 10 Gaps: 10



```
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db 2224 ATTGATGATCTCTGAACAAACACGACAGAAAGTAGTAACCTTTTTCACATATTGTTCTTAGGCC 2283
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db 2284 CTTGGGAATTATTTCTTTTATTACATTTTCTTCAAGGTTTTCATTTGGCAAGCTGGA 2343
QY 727 GluIleLeuMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db 2344 GAGATCTCTACCAAGCGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTG 2403
QY 747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db 2404 AGTTGGTTTGTATGATCCCTAAACACACACCTGGAGCATTTGACTACCGAGCTCGCCAAATGAT 2463
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2464 GCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTCTGTAATTATCCAGAAATATAGCA 2523
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db 2524 AATCTTGGGACAGGAATAATATATATCTTCACTATGTTGGCACTTAACACTGTTTACTC 2583
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 2584 TTAGCAATTGTACCATTCATTCGCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTCTGGA 2643
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAAGTGCTGGGAAGATCGCTACTGAAGCAATA 2703
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 2704 GAAACATTCGGAACCGTGTCTTTCGACTCAGGAGGAAGTTTGAACATATGTTATGCT 2763
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 2764 CAGAGTTTGAGGTTGACCATACATACAGAAACTCTTTGAGGAAGACACACATCTTTGGAATTACA 2823
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 2824 TTTTCTTCAACCCAGCAATGATGATTTTCTTCTATGCTGATGTTTCCGGTTTGGAGCC 2883
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 2884 TACTTGTGGCACATAAATCATCAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTT 2943
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 2944 GTCTTTGGTGCATGGCGTGGGCAAGTCAGTTCAITTTGCTCTGACTATGCCAAAGCC 3003
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArg 966
Db 3004 AAAATATCAGCAGGCCACATCATCATGATCATTTGAAAAAACCCTTTGATTTGACAGCTAC 3063
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3064 AGCA CGGAAGCCCTAATGCCAAACATCTGGAAGGAAATGTCATTTGGTGAAGTTGTA 3123
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3124 TTCAACTATCCACCCGACCCGACATCCAGTGTCTTCAGGACTGAGCCTGGAGGTGAAG 3183
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3184 AAGGCCACAGCGCTGCTGCTGGGACGAGTGGCTGTGGGAAGAGACACAGTGGTCCAG 3243
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db 3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGAAAGTGTGCTTGTATGGCAAGAAATA 3303
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
```

```
Db 3304 AAGCGACTGAATGTTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATC 3363
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3364 CTTGTTGACTGACGATTCCTGAGAACATTTGCTATGGAGACAACAGCCGGGTGGTGTCA 3423
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3424 CAGGAAGAGATCGTAGGGCAGCAAAAGGAGGCCAACATACATGCTTTCATCGAGTCACTG 3483
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLys 1126
Db 3484 CCTAATAAATATAGCACTAAAGTAGGAGCAAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 3544 CAACGATTTGCCATAGCTCGTGCCTTTGTAGACACCTCATATTTTGTCTTTTGGATGAA 3603
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3604 GCCACGTCAGCTCTGCATACAGAAAGTGAAAGGTTGTCCAAGAACCCCTGGACAAAGCC 3663
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3664 AGAAGAAGCGCGCACCTGCTGCTGTTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGAC 3723
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 3724 TTAATAGTGTGTTTTCAGAAATGCGCAGATGCGCAGGACATGGCAGCATCAGCAGTGTG 3783
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 8
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:
Pred. No.:          9,87e-313      Length:      8630
Score:              3521.50        Matches:    681
Percent Similarity: 73.92%         Conservative: 246
Best Local Similarity: 54.31%      Mismatches:  264
Query Match:       57.18%          Indels:      63
DB:                10              Gaps:       10

US-09-873-409-5 (1-1222) x US-09-306-417-1 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AenLeuIleSer 31
DB 1430 GTGTTTGGAGAAATGACAGATATCTTTGCCAAATGCAGGAATTTAGNAGATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AenThrTySerPhePhe----- 43
DB 1490 AACATCACTAATAAGAGTGCATATCATGATACAGGGTCTCTTCATGAATCTGGAGGAAGAC 1549
QY 44 -----ArgLeuThrLeuTyThrValGlyIleGlyValAlaAlaLeuIlePheGlyTy 61
DB 1550 ATGACCAGATATCCCTATTATTACAGTGAATTTGGTCTGGGTGCTGGTCTGCTTAC 1609
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleArgIleGln 81
DB 1610 ATTCAGGTTTCATTTTGGTCCCTGGCAGCTGGAGACAAATACACAAATTAGAANAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATAATCGCAGCAGAGATAGGCTGGTTTGTATGCACGATGTTGGGAG 1729
QY 102 LeuAsnThrArgMetThr----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAAACCCCGACTTACAGATGATGTCTCTAAGATTAAATGAAGTTATTGGTGACAAAATT 1789
QY 121 AlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTGTGAAGCTAACCCCTGTGATTTTGGCCATCAGTCTCTTCTTGAGCTGCTGCTGCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTySerLysAla 180
DB 1910 GTCTGGGCAAGATACATATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1970 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGTGATTCGATTTGGAGACAA 2029
QY 201 GlyLysGlnLeuGlnArgSerPheLeuLeuAsnIleThrArgTyThrAlaTrpPheTyPhe 220
DB 2030 AAGAAAGAACTTGAA----- 2044
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyThrGlnAenLeuLysAsp 240
DB 2045 -----AGGTACAACAAAATTTAGAAAGAA 2068
```





FEATURE: OTHER INFORMATION: Description of Artificial Sequence: proviral  
OTHER INFORMATION: plasmid DNA  
FEATURE: NAME/KEY: misc feature  
LOCATION: (1)..(8630)  
OTHER INFORMATION: retroviral expression vector SFbeta91msA1  
FEATURE: NAME/KEY: misc feature  
LOCATION: (1)..(160)  
OTHER INFORMATION: plasmid backbone (pUC)  
FEATURE: NAME/KEY: misc feature  
LOCATION: (161)..(677)  
OTHER INFORMATION: 5'-LTR  
FEATURE: NAME/KEY: 5'UTR  
LOCATION: (532)..(1219)  
FEATURE: NAME/KEY: mat\_peptide  
LOCATION: (1220)..(5062)  
OTHER INFORMATION: msA1 mdr1 cDNA  
FEATURE: NAME/KEY: misc feature  
LOCATION: (5215)..(5774)  
OTHER INFORMATION: 3'-LTR  
FEATURE: NAME/KEY: misc feature  
LOCATION: (5775)..(8630)  
OTHER INFORMATION: plasmid backbone (pUC)  
US-09-306-417-2

Alignment Scores:  
Pred. No.: 9,87e-313 Length: 8630  
Score: 3521.50 Matches: 681  
Percent Similarity: 73.92% Conservative: 246  
Best Local Similarity: 54.31% Mismatches: 264  
Query Match: 57.18% Indels: 63  
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-306-417-2 (1-8630)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
|||.....|
Db 1370 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATGCTG 1429
|||.....|
QY 21 ValLeuGlyGluMetSerAsp-----AnLeuIleSer 31
|||.....|
Db 1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGAAATTTAGAAAGATCTGATGTCA 1489
|||.....|
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
|||.....|
Db 1490 AACATCACTAATAGAGTATCAATGATACAGGGTTCCTGGAATCTGGAGGAAGAC 1549
|||.....|
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
|||.....|
Db 1550 ATGACCAAGGTATGCCATTATTACAGTGAATTTGGTCTGGGGTGGTGGTGGCTGTAC 1609
|||.....|
QY 62 IleGlnIleSerLeuThrIleLeuAlaArgGlnThrLysArgIleArgLysGln 81
|||.....|
Db 1610 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 1669
|||.....|
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
|||.....|
Db 1670 TTTTTCATGCTATAATGGCAGAGATAGGCTGGTTTATGATGTCACGATGTGGGGAG 1729
|||.....|
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
|||.....|
Db 1730 CTTAAACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTTGGTACAAATTT 1789
|||.....|
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
|||.....|
Db 1790 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849
|||.....|
```

```
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
|||.....|
Db 1850 GGTTCGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTGTCTTGGACTGTGCAGTGTCT 1909
|||.....|
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
|||.....|
Db 1910 GTCTGGCAAGATATCTATCTTCACTTATTAAGAACTCTTACGCTAAGCAAGCT 1969
|||.....|
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
|||.....|
Db 1970 GAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAATGTGATTCATTTGGAGGACAA 2029
|||.....|
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
|||.....|
Db 2030 AAGAAAGAACTTGAA----- 2044
|||.....|
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
|||.....|
Db 2045 -----AGGTACACAAATAATTTAGAGAA 2068
|||.....|
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
|||.....|
Db 2069 GCTAAAGAAATTTGGATAAAGAAAGCTATTACAGCCAATATTCTATAGTGTCTGCTTTC 2128
|||.....|
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
|||.....|
Db 2129 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGTGATGGGACCACTTGTGCTCTCTCA 2188
|||.....|
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
|||.....|
Db 2189 GGGGAA-----TATTCTATTGGACAAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 2242
|||.....|
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
|||.....|
Db 2243 GCTTTTAGTGTGGACAGCACTCCAGCATTTGAAGCAATTTGCCAATGCAAGAGGAGCA 2302
|||.....|
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
|||.....|
Db 2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAATATTGACAGCTATTGCAAGAGT 2362
|||.....|
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
|||.....|
Db 2363 GGGCACAACCCAGATATATTAAAGGAAATTTGGAATTTGCAAAATGTTCATCTCAGTTAC 2422
|||.....|
QY 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
|||.....|
Db 2423 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGCTGAACTGAAGTGCAGATGGGCGAG 2482
|||.....|
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
|||.....|
Db 2483 ACGTGGCCCTGTTGGAAACAGTGGCTGGGGAAGACACACAGTCCAGCTGATGTCAG 2542
|||.....|
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
|||.....|
Db 2543 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACCAT 2602
|||.....|
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
|||.....|
Db 2603 AATGTAGGTTTCTACGGGAATCATTTGGTGTGTGAGTCAAGAACTGTATTGTTTGGC 2662
|||.....|
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
|||.....|
Db 2663 ACCACGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCCACCTGATGATGATGAG 2722
|||.....|
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
|||.....|
Db 2723 AAAGCTGTCAAGGAAGCCAAATGCTATGACTTTATCATGAACTGCCTCATAAATTTGAC 2782
|||.....|
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
|||.....|
Db 2783 ACCCTGTTGGAGAGAGGGGCCCACTTGTAGTGGTGGGAGAGCAGGATGCGCCATT 2842
|||.....|
```





```

Db 5003 GCACAGAAAGCATCTATATTTTTCATGCTCAGTGTGTCAGGCT 5044

RESULT 10
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.: 9,44e-313 Length: 4425
Score: 3517.00 Matches: 680
Percent Similarity: 73.90% Conservative: 243
Best Local Similarity: 54.44% Mismatches: 266
Query Match: 57.10% Indels: 60
DB: 10 Gaps: 10

US-09-873-409-5 (1-1222) x US-09-769-097-3 (1-4425)

Qy 1 MetlleuGlyVlleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 499 ATGCTGTGGGAACCTCTGGCGGCCATTATCCATGGAATTGGCTGCCATTATGATGCTG 558
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 559 GTCTTTGGAGACATGACAGATAGCTTTTGCAAT-----GTAGGAAACACCGTAGTATG 612
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 41 SerPhePheArgLeuThr----- 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 613 AGTTTCTACAAATGCTACAGACATATATGCCAAGCTGGAGACGAAATGACCAGTAGGCC 672
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 47 LeuTyrTyrValGlyVlleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 673 TACTATTACACGGGCATTGGTCCGGTGTCTCATCGTTCCTACATCCAGGTTTCCACT 732
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 733 TGGTGCCTGGCAGATGGGAGACAAATACAAAGATTAGGCAGAAAGTTTTTCCATGCCATC 792
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 87 LeuAlaGlnAspIleGlyTrrPheAspSerCysAspIleGlyGluLeuAenThrArgMet 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 793 ATGAATCAGGAGATAGCGTGTGTTGACGTGCATGACGTTGGGAGCTCAACACCCGGCTC 852
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 853 ACAGATGAGCTCTCCAAATTAATGAAGAAATTCGTGACAAATTGGAAATGTTCTTTCAG 912
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrrLysLeuThr 145
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 913 GCMAATGGCAACATTTTGGTGGTTTTATATAGCATTTACTCGCGCTCGAAGCTAACT 972
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

|||||  
Db 1966 CGCAACCCCAAGATCTTTTGGATGAGCCACGTCAGCCTTGGACACAGAAAGCGAA 2025  
Qy 526 SerAlaValGlnAlaLeuGluLeuAlaSerLysGlyArgThrThrIleValValAla 545  
Db 2026 GCCGTGGTTCAGGCGCTCTGGATAGGCTAGAGAGCGGACCCACCATTTGATGAGCT 2085  
Qy 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565  
Db 2086 CACCGCTTGTCTACAGTTCGCAATGCTGATGTCATTCCTGGTTTGGTGGTGCATT 2145  
Qy 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrThrSerLeuVal 585  
Db 2146 GTGAGCAAGGAATCATGATGAGCTCATGAGAGAGAAGGAATTTACTTCAAACTTGTG 2205  
Qy 586 MetSerGln-----AspIleLysLysAlaAspGluGlnMetGluSer----- 599  
Db 2206 ATGACTCAGACAGCAGGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 2265  
Qy 600 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeu 613  
Db 2266 GGAATTGATATGTGGACATGCTTCAAGAGATTCRGATCCAGTCTAATAAGAGAGA 2325  
Qy 614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632  
Db 2326 TCAACTCGCAAAAGCATCGTGGGCCACATGATCAAGACGGGGAACCTTAGCACCAGAG 2385  
Qy 633 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 652  
Db 2386 GCTCTGATGACGACGACCTCCAGCTCCCTTTGGCGGATCCCTGAAGTTGAATTCAACT 2445  
Qy 653 GluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672  
Db 2446 GAATGGCCCTTATTTTGGTGGTGTATTTGTGCCATATAAATAGAGCGCTTGCACCA 2505  
Qy 673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThr 692  
Db 2506 GCATTCCTCATATATTTCAAGAGTTGTAGGGGTTTTTCAAAAAAATGACACCCCTGAA 2565  
Qy 693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711  
Db 2566 ATCCAGGGCAGAACAGCACTTTTCTTATTTATTTCTTCATCTTCCTGGGATCATCTCT 2625  
Qy 712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731  
Db 2626 TTCATTAGTTTTTCTTCAGGGCTTACATTTGGCAAGAGCTGGAGAGATCTCACCAG 2685  
Qy 732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751  
Db 2686 CGACTCCGATACATGCTTCAAAATCCATGCTGAGACAGGACATAAGCTGGTTTGTATGAC 2745  
Qy 752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771  
Db 2746 CCTAAACACACAGGAGGCGCTGACACACAGGCTTGGCAATGACGCTGCTCAAGTAAA 2805  
Qy 772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791  
Db 2806 GGGCTACAGGGCTTAGGCTTGTGTTATACCCAGAACATAGCAATCTTGGACAGGC 2865  
Qy 792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811  
Db 2866 ATCATCATATCCCTGATCTACGGCTGGCAATGATGACATTTTACTCTTAGCAATTTGCC 2925  
Qy 812 ValLeuAlaValThrGlyMetIleGluThrAlaIleMetThrGlyPheAlaAsnLysAsp 831  
Db 2926 ATCATGTATAGCAGGAGTGGTTGAAATGAAATGTTTGTCTGCAAGCGCTGAAAGAT 2985  
Qy 832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851  
Db 2986 AAGAGGAACTTAGAAGGTTCTGGGAGATCGCTACAGAGCAATGAATTAATTTTCGACT 3045  
Qy 852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871  
Db 3046 GTCGTCTCTTTGACTCGGAGCGAAGAGTTTGAAACTATGTATGATCCCGAGAGCTTGCAGATA 3105

Qy 872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891  
Db 3106 CCATACAGAAATGCTTTGAGAAAGCGCAGCTCTTTGGGATCACTTTCTCTTCAACCAG 3165  
Qy 892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911  
Db 3166 GCCATGATGATTTCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3225  
Qy 912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931  
Db 3226 GAATCATGACATTTGAAATGTTCTGTAGTATTCTAGCTATTGTCTTTGTGGCCATG 3285  
Qy 932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951  
Db 3286 GCAGTGGGCGAGTCAGTTCATTCGCTCTCTGACTACGCGAAACCAAGTCTCGGCATCC 3345  
Qy 952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971  
Db 3346 CACATCATCAGGATCATTTGAGAAATCCCTGAGATTGACAGCTACAGCAGCGAGGCTTG 3405  
Qy 972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991  
Db 3406 AAGCCTAATATGTTGAAAGGAATGTAAATTTAATGGAGTCATGTTCAACTATCCACC 3465  
Qy 992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011  
Db 3466 CGACCCCAACATCCAGTGTCTCAGGGGCTGAGCTAGAGGTGAAGAAAGGCGACGCTG 3525  
Qy 1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031  
Db 3526 GCCTCTGGGCGACGAGTGGCTCGGGAAGAGTACAGTGGTCCAGCTGCTGTGAGCGCTTC 3585  
Qy 1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051  
Db 3586 TATGACCCCATGCCGGAACAGATGTTTCTAGATGGCAAGAAATAAGCAACTCAATGTC 3645  
Qy 1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071  
Db 3646 CAGTGGCTCGCGCCACCTGGGCAATGTGTCCAGGAGGCCATCTCTGTTGATGTCAGC 3705  
Qy 1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091  
Db 3706 ATCCGCGAGAACTTGCCTACGGAGACACAGCGGTGCTGTCTCATAGGAGATCGTG 3765  
Qy 1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111  
Db 3766 AAGGCAAGCCAGGAGGCAACATCCACAGTTCATCGACTCCTGCTGAGAAATACAC 3825  
Qy 1112 ThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131  
Db 3826 ACAGAGTGGAGACAAAGGACTCAGTCTCGGCGGGCAGAGCAGCGCATCGCCATC 3885  
Qy 1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151  
Db 3886 GCGCGCGCCCTCGTCAGACAGCCTCATCTTACTTCTGATGAAGCGACATCAGCTCTG 3945  
Qy 1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyValThr 1171  
Db 3946 GATACGAGAGTGAAGAGTGGTCCAGAGAGCGCTGGACAAAGCCAGGAGGCGCGCAC 4005  
Qy 1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191  
Db 4006 TGCATTGTGATCGGCACCGCCTGTCCACCATCCAGAACCGCAGACTTGTGCTGGTGATT 4065  
Qy 1192 HisAsnGlyLysLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211  
Db 4066 CAGAAACGCGCAGGTCAAGGAGCAGCGCACCCACAGCAGCTGCTGGCCCGAAGAGGCATC 4125  
Qy 1212 TyrPheLysLeuValAsnAlaGlnSer 1220  
Db 4126 TATTTCTCGATGGTCACTGTGCAAGCT 4152

## RESULT 11

```
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:      1,41e-312      Length:      4369
Score:          3515.00      Matches:      680
Percent Similarity: 73.90%      Conservative: 243
Best Local Similarity: 54.44%      Mismatches:  266
Query Match:      57.07%      Indels:       60
DB:                10          Gaps:         10

US-09-873-409-5 (1-1222) x US-09-769-097-1 (1-4369)
QY 1 MetIleuGlyIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 499 ATGCTGTGGGAACCTCTGGCGGCATTATCCATGGAAATCGCTCCACATTATGATGCTG 558
QY 21 ValLeuGlyGluMetSerAspAsnLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAAT-----GTAGAAACAACCGTAGTATG 612
QY 41 SerPhePheArgLeuThr----- 46
DB 613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGGCCACGTACGCC 672
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 673 TACTATTACACGGGCAATGGTGGCGGTGCTCATCGTTCATCGCTACATCCAGGTTTCATT 732
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
DB 733 TGGTGCCTGGCAGATGGGTGGTGTGACGTGCATGACCTAGGAGGAGTGTTCATGCCCATC 792
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 793 ATGAATCAGGAGATAGGCTGGTGTGACGTGCATGACCTTGGGGAGCTCAACACCCGCGCTC 852
QY 107 Thr---AspIleAspIlySerAspGlyIleGlyAspIlySerIleAlaLeuLeuPheGln 125
DB 853 ACAGATGACGCTCTCAAAATTAATGAAGGAATGGTGACAAAATTGGAATGTTCTTTTCAG 912
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrPheLeuThr 145
DB 913 GCAATGGCAACATTTTTTGGTGGTTTTATATAGGATTTACTCGCGGCTGGAAGCTAACT 972
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
```

```
Db 973 CTTGTGATTTGGCCATCAGCCCTGTCTTGGACTGTCAGCTGGTATTTGGGCAAGATA 1032
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db 1033 TTGTCTTCAATTTACTGATGAAGAACTCCAGGCTTATGCAAAAGCTGAGCAGTTGCTGAA 1092
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db 1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCTTTGGAGGACAAAGAAGAACTTGAA 1152
QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeu 225
Db 1152 ----- 1152
QY 226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db 1153 -----AGGTACAATAACAAATTTGGAAGAAGCTAAAGAGCCTGGG 1191
QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db 1192 ATAAAGAAAGCTATCACGGCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA 1251
QY 266 ThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db 1252 TCATATGCTCTGGCAATTCGTATGGGACTTCCTTGGTCAATCTCAAAAGAA-----TAC 1305
QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db 1306 ACTATTGGACAAGTGTCTCACTGCTTTTTTTTCTGTAATTAATGGAGCATTCACTGTTGGG 1365
QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db 1366 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTAAGTCTTC 1425
QY 326 GlnValIleAspLysIlySerProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db 1426 AGTATAATTGATAAATAAGCCAGTATAGACGCTTCTCAAAAGAGTGGGCACAAACCCGAC 1485
QY 346 SerIleGlyGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db 1486 ACATACAGAAATTTGGAATTCAAANAATATTCATCTCAGTTACCCGCTCGAAAGAC 1545
QY 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db 1546 GTTCAGATCTTGAAGGCCTCAACCTGAAGGTGAAGCGGCGAGCGTAGCCCTGGTT 1605
QY 386 GlyLeuAsnGlySerGlyLysSerThrValGlnIleLeuGlnArgLeuTyrAspPro 405
Db 1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGTGCAGAGGCTCTACGACCCC 1665
QY 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db 1666 ATAGAGGGCGAGGTAGTATFCGACGGAGGAGGACCATCAGGACCATCAATGTGAGGTATCTG 1725
QY 426 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 445
Db 1726 CGGGAATCATTTGGGTGGTGTGAGTCAGGAACCGCTGCTGTGTTCGCCACCAATTCGCGAA 1785
QY 446 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db 1786 AACATTTCGCTATGGCGAGAAAACGTCACCATGGATGAGTAGAGAAGCTGTCAAGGAA 1845
QY 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnIlyPheAsnThrLeuValGlyGlu 485
Db 1846 GCCAATGCCCTATGATTTCAATCATGAATCCGCCCAAAATTTGACACCTTGGTGGTGAG 1905
QY 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db 1906 AGAGGGCGGAGCTGAGTGGGGACAGAAACAGAGGATCGCCATTCGCCCGGCGCTGTC 1965
QY 506 ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
```

Db 1966 CGAACCCCAAGATCCTTTTGTGGATGAGCCACAGTCAGCCTTGGACACAGAAAGCGAA 2025  
Qy 526 SerAlaValGlnAlaAlaIleGluIleValSerIleGlyArgThrThrIleValAla 545  
Db 2026 GCCGTGTTGAGCGCTCTGGATGAGGCTAGAGAGCGGACCCATTTGATGACT 2085  
Qy 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuIleAspGlyMetLeu 565  
Db 2086 CACCGCTGTCTACAGTTGCAATGCTGACGCTCATCTGCTGTTGATGGTGTCATT 2145  
Qy 566 AlaGluIleGlyAlaHisAlaGluLeuMetAlaIleArgGlyLeuIleThrSerLeuVal 585  
Db 2146 GTGAGAGCAAGAAATCATGATGACTCATGAGAGAGAAAGAAATTTACTTCAAACTTGTC 2205  
Qy 586 MetSerGln-----AspIleIleGlyAlaAspGluGlnMetGluSer----- 599  
Db 2206 ATGACTCAGACAGCAAGAAATGAATTTGAATTTAGAAATGAAGCTTGTGAATCTTAAGAC 2265  
Qy 600 -----MetThrTyrSerThrGluArgIleGlyThrAsnSerLeu-----ProLeu 613  
Db 2266 GGAATTGATGATGGACATGCTCTTCAAAAGATTTCGGGATCCAGTCTTAATAAGAAGA 2325  
Qy 614 HisSerValIleSerIleIleSerAspPheIleAspIleValGluGluSerThr---Gln 632  
Db 2326 TCAACTCGCAAAAGCATCCGTGGCCACATGATCAAGACGCGGAACTTAGCACCAGAG 2385  
Qy 633 SerIleGluIleSerLeuProGluValSerLeuIleIleIleIleIleIleIleIleIle 652  
Db 2386 GCTCTGATGACGACGATCCAGCTTCTTTTGGCGGATCTTGAAGTTGAATTCAACT 2445  
Qy 653 GluTrpProPheValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672  
Db 2446 GAATGGCTTATTTTGGTGGTGTATTTGTGCCATTAATTAATGAGGCTTGCAACCA 2505  
Qy 673 ValPheSerIleIlePheAlaIleIleIleThrMetPheGlyAsnAsnAspIleThr 692  
Db 2506 GCATTCCTCAATAATTTTCAAGAGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAA 2565  
Qy 693 Leu---IleHisAspAlaGluIleIleSerMetIlePheValIleLeuGlyValIleCys 711  
Db 2566 ATCCAGCGGCAACAGCACTTTTCTTTATTTGTTCTGATCTTGGGATCATCTCT 2625  
Qy 712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731  
Db 2626 TTCAATAGCTTTTCTTCAGGCTTCCATTTGCGCAAGCTGGAGAGATCTCCACCAAG 2685  
Qy 732 ArgLeuArgHisLeuAlaPheLeuAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751  
Db 2686 CGACTCCGATACATGCTCTTCAAAATCCATGCTGACAGCAGGACATAAGCTGTTGATGAC 2745  
Qy 752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771  
Db 2746 CCTAAAAACACCAAGAGCGCTGACCAAGGCTTCCCAATGACGCTGCTCAAGTGAAA 2805  
Qy 772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791  
Db 2806 GGGCTACAGGCTCTAGGCTGCTGTTATTTACCCAGAACATAGCAATCTTGGACAGGC 2865  
Qy 792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811  
Db 2866 ATCATCATATCCCTGATCTACGGCTGGCAATTGACATTTTACTCTTAGCAATTTGTTCCC 2925  
Qy 812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831  
Db 2926 ATCAATGCTATAGCAGGAGTGGTTGAATTTGAATTTGTTCTGGACAGCGCTCAAGAT 2985  
Qy 832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851  
Db 2986 AAGAGGAACTAGAGGTTCTGGGAAGATCCCTACAGAGCAATTTGAATTTTCGACT 3045  
Qy 852 IleValSerLeuThrArgGluIleAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871  
Db 3046 GTCTGCTCTTTGACTCGGAGCAGAAAGTTTGAACATGATGATGCCAGAGCTTGACGATA 3105

Qy 872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891  
Db 3106 CCATACAGAAATCCTTTGAAGAAAGCGACGCTTTGGGATCATCTTCTCTTACCAC 3165  
Qy 892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911  
Db 3166 GCCATGATGATTTCTCTCTATGCTGCTGTTCCGGTTCATGCTTACTTGTGGTGCACGA 3225  
Qy 912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931  
Db 3226 GAATCATGACATTTGAAATGTTCTGTAGTATTCTCAGCTATTGTTTGTGGCATG 3285  
Qy 932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951  
Db 3286 GCAGTGGGCGAGCTCAGTTCATTCGCTCTCTGACTACGCGAAAGCAAGTCTCGGCATCC 3345  
Qy 952 HisLeuPheAlaLeuLeuGluLysIleProAsnIleAspSerArgSerGlnGluGlyLys 971  
Db 3346 CACATCATCAGGATCATTTGAGAAATCCCTGAGATTGACAGCTTACAGCAGCGAGGCTTG 3405  
Qy 972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991  
Db 3406 AAGCCTAATATGTTGAAGGAAATGTAATTTAATGGAGTCAATGTTCAACTATCCACC 3465  
Qy 992 ArgProAspValPheIleLeuArgGlyLeuSerLeuIleGluArgGlyLysThrVal 1011  
Db 3466 CGACCCCAACATCCAGTGTCTCAGGGCTGAGCTAGAGGTGAAGAAAGGCGACGCTG 3525  
Qy 1012 AlaPheValGlySerSerGlyCysGlyIleSerThrSerValGlnLeuLeuGlnArgLeu 1031  
Db 3526 GCCTCGTGGGACGAGTGGCTCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTC 3585  
Qy 1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051  
Db 3586 TATGACCCCATGCGCGAAGCAGTGTCTTAGTGGCAAGAAATAAGCAACTCAATGTC 3645  
Qy 1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071  
Db 3646 CAGTGGCTCGCGCCCACTGGGCATTTGTCCAGAGGCCCATCTGTTGATGTCAGC 3705  
Qy 1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091  
Db 3706 ATCGCGAGAAATTCCTACGAGACCAACAGCCGTGTCGTGTCATAGAGGAGATCGTG 3765  
Qy 1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluClyLeuProGluLysTyrAsn 1111  
Db 3766 AAGCGACCCCAAGAGGCGCAACATCCAGTTCATCGACTCATGCTGCTGAGAAATACAAC 3825  
Qy 1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131  
Db 3826 ACCAGTGGGACAAAGGACTCAGCTGTCGGCGGGCAGAGCAGCGCATCGCATC 3885  
Qy 1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151  
Db 3886 GCGCGCGCTCTCGTCAGACAGCTCATCTTACTTCTGGATGAAGCGCATCAGCTCTG 3945  
Qy 1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171  
Db 3946 GATACGGAGAGTGAAGAGTGTCCAGGAGCGCTGGACAAAGCCAGGGAAGCGCGCACC 4005  
Qy 1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191  
Db 4006 TGATTTGTGATCGGCACCGCTGTCCACCATCCAGAACCGACACTTGTGTTGATG 4065  
Qy 1192 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211  
Db 4066 CAGAACCGCGCAGGTCAAGAGACCGGCACCCACAGCAGCTGCTGGCCCAAGAGGCATC 4125  
Qy 1212 TyrPheLysLeuValAsnAlaGlnSer 1220  
Db 4126 TATTTCTGATGTCAGTGTGCGAGGCT 4152

## RESULT 12

US-10-072-621-2  
; Sequence 2, Application US/10072621  
; Patent No. US20020169137A1  
; GENERAL INFORMATION:  
; APPLICANT: Reiner, Peter B.  
; APPLICANT: Connop, Bruce P.  
; APPLICANT: Pollard, Michelle  
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION  
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY  
; FILE REFERENCE: 100103.402  
; CURRENT APPLICATION NUMBER: US/10/072,621  
; CURRENT FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 4643  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-072-621-2

Alignment Scores:  
Pred. No.: 1,916-312 Length: 4643  
Score: 3514.00 Matches: 681  
Percent Similarity: 73.92% Conservative: 246  
Best Local Similarity: 54.31% Mismatches: 263  
Query Match: 57.05% Indels: 64  
DB: 9 Gaps: 11

US-09-873-409-5 (1-1222) x US-10-072-621-2 (1-4643)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 575 ATGGTGGTGGGAACCTTTGGCTGCATCCATCCATGGGCTGGACATCTCTCATGATGCTG 634  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAATGCAGGAAATTTAGAAATCTGATGTCA 694  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 695 AACATCACTAATAGAAATGATATATGATGATACAGGGTTCCTTCATGATCTGGAGGAGAC 754  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 755 ATGACACAGGTATCCCTATTATTACAGTGGAAATTTGGTCTGGGTGCTGGTGTCTTAC 814  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 815 ATTCAGGTTTCATTTTGGTCCCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTTCATGTCACGATGTTGGGAG 934  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 935 CTTAACACCCGACTTACAGATGATGCTCCAGATTAATGAGGAATTTGGTGACAAATTT 994  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 995 GGAATGTTCTTCAGTCAATGGCAACATTTTTCATCGGTTTATAGTAGATTATACAGT 1054  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 1055 GGTGGGAAGTAACCTTTGTGATTTTGGCCATCATGCTCTCTTCTTGGACATGTCAGCTGCT 1114  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180  
DB 1115 GTCTGGCAAGATATCTTCTTATTACTGATTAAGAACTCTTAGCGTATGCAAGCT 1174  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

DB 1175 GGAGCAGTAGCTGAAGAGGCTCTTGGCAGCAATTAGAACTGTGATTGTCATTTGGAGACAA 1234  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 1235 AAGAAAGAACTTGA----- 1249  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 1250 -----AGGTACACAAATAATTAGAGAA 1273  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 1274 GCTAAAGAAATTTGGATAAAGAACTATTACAGCAATATTCTATAGTGTGCTGCTTTC 1333  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsn 280  
DB 1334 CTGCTGATCTATGATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGTGCTCTCA 1393  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 1394 GGGGAA-----TATTTCTATTGGCAAGTACTCAGTGA---TTTTCTGTATTAATTGGG 1444  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 1445 GCTTTTATGTTGGACAGGATCTCCAGCATTGAAGCATTTCGAAATGCAAGAGAGCA 1504  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 1505 GCTTATGAATCTTCAAGATAATTGATAAAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1564  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 1565 GGGCACAACACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTTCAGTTAC 1624  
QY 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1625 CCATCTCGAAAAAGATTTAAGATCTTGAAGGGTCTGAACCTGAAGGTGACAGTGGGCG 1684  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1685 ACGTGGCCCTGGTTGGAAACAGTGGCTGGGAAGAGACACACAGTCCAGCTGATGCGAG 1744  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
DB 1745 AGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACATA 1804  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1805 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGGTGCAGTCAGGAACCTGTTATTTTGC 1864  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 1865 ACCACGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCCACCATGGATGAGATTGAG 1924  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1925 AAGCTGTCAAGGAAGCAATGCTATGATTTTATCATGAAACTGCTCATATAATTGAC 1984  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1985 ACCCTGGTTGGAGAGAGAGGGCCAGTGTGAGTGGGGCAGAGCAGAGATTCGCCATT 2044  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 2045 GCACGTGCCCTGGTTGCGAACCCCAAGATCTCTCTGCTGGATGAGGCCAGCTCAGCCTTG 2104  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
DB 2105 GACACAGAAAGCGAAGCAGTGGTTTCAAGTGGCTCTCGATAAGGCCAGAAAGGTCGAGCC 2164  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
DB 2165 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTTCTGTAATGCTGACGTATCGCTGCTTTC 2224



QY	561	LyeAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu	580
DB	2225	GATGATGGAGTCATTCTGGAGAAAGAAATCATGATGAATCATGAAAGAGAAAGGCATT	2284
QY	581	TyTyrSerLeuValMetSerGln-----AspLleLysLysAlaAsp	594
DB	2285	TACTTCAAACTTTGTCACAAATGCAGACAGCAGGAAATGAAGTTGAAATAGAAAAATGCAGCT	2344
QY	595	GluGlnMetCysSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis	614
DB	2345	GATGAATCCAAAAGTGAATAATGATCCCTTGGAAATGTCCTTCAAAATGATTCAGATCCAGT	2404
QY	615	SerVal-----LysSerLleLys---SerAspPheLleAspLysAla	627
DB	2405	CTAATTAAGAAAAGATCAACTCGTAGGAGTGTCCGTGATCACAAGCCCAAGACAGAAAG	2464
QY	628	GluGluSerThrGlnSerLysGluLleSerLeuProGluValSerLeuLeuLysLleLeu	647
DB	2465	CTTAGTAGCAAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTGGAGGATTATG	2524
QY	648	LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn	667
DB	2525	AAGCTAAATTTAACTGAATGGGCTTATTTTGTGTGTGTATTTTGTGGCATTTATAAAT	2584
QY	668	GlyThrValHisProValPheSerLleLlePheAlaLysLleLleThrMetPheGlyAsn	687
DB	2585	GGAGGCTGCACACGACATTCATATATATTTTCAAAGATATATAGGGTTTTTACAAGA	2644
QY	688	---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetLlePheValIle	706
DB	2645	ATTGATGATCCTGAAACAAACACAGAGAATAGTAACCTTGTTTTCACTATTGTTTCTAGCC	2704
QY	707	LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly	726
DB	2705	CTTGGAAATATTTCCTTTTAAATACATTTTCTTCAGGGTTTTCACATTTGGCAAGCTGGA	2764
QY	727	GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle	746
DB	2765	GAGATCCTCACCAAGCGCTCCGATACATGGTTTCCGATCCATGCTCAGACAGGATGTG	2824
QY	747	AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp	766
DB	2825	AGTTGGTTTGATGACCTTAAACACACCACCTGGAGCATTTGACTACCAGGCTCGCCAAATGAT	2884
QY	767	IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr	786
DB	2885	GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTTACCCAGAATATAGCA	2944
QY	787	AsnMetClyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle	806
DB	2945	AACTCTGGGACAGAAATAATATATCCTTCATCTATGGTTGGCACTTAACACTGTTACT	3004
QY	807	LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly	826
DB	3005	TTAGCAATTGTCACCATCATTCGAATAGCAGAGTTGTTGAAATGAAAAATGTTGCTGGA	3064
QY	827	PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu	846
DB	3065	CAAGCACTGAAGACATAAGAAAGAACTAGAAAGGTGCTGGAGAGATCGCTACTGAAGCAATA	3124
QY	847	GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu	866
DB	3125	GAAAACTTCCGAACCGTTGTTCTTTTGACTCCAGGACGAGAGTTTGAACATATGTATGCT	3184
QY	867	GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleClySerCys	886
DB	3185	CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGACACATCTTTTGGAAATTACA	3244
QY	887	TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla	906
DB	3245	TTTTTCTTCAACCAGGCAATGATGATATTTTCTATGCTGGATGTTTCCGGTTTGGAGCC	3304

RESULT 13

US-09-866-866A-5

US-09-000-000A-3  
; Sequence 5, Application US/09866866A

; Patent No. US20020102244A1

; FACEID NO: 032002010  
; GENERAL INFORMATION:

APPLICANT: Sorrentino, Brian

: APPLICANT: Schuetz, John

APPLICANT: SCHUEZ, JOHNNIE  
TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

FILE REFERENCE: 1340-1-021CIP2



; CURRENT APPLICATION NUMBER: US/09/866, 866A  
 ; CURRENT FILING DATE: 2001-08-30  
 ; PRIOR APPLICATION NUMBER: 09/584, 586  
 ; PRIOR FILING DATE: 2000-05-31  
 ; PRIOR APPLICATION NUMBER: PCT/US99/11825  
 ; PRIOR FILING DATE: 1998-05-27  
 ; PRIOR APPLICATION NUMBER: 60/086, 988  
 ; PRIOR FILING DATE: 1998-05-28  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 5  
 ; LENGTH: 4189  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 US-09-866-866A-5

Alignment Scores:  
 Pred. No.: 1,68e-311 Length: 4189  
 Score: 3503.00 Matches: 678  
 Percent Similarity: 73.13% Conservative: 242  
 Best Local Similarity: 53.90% Mismatches: 262  
 Query Match: 56.88% Indels: 76  
 DB: 10 Gaps: 11

US-09-873-409-5 (1-1222) x US-09-866-866A-5 (1-4189)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 148 ATGATTCTGGGAACCTCTCGCTGCTATATCCATGGAAACATTACTTCCCTCTTGATGCTG 207  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30  
 DB 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTAGTTCTGCTCAAGCAATT 267  
 QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
 DB 268 ACTAATCAAGTGGACCCACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327  
 QY 40 -----TySerPhePheArgLeuThrLeuTyTyValGlyIleGlyValAlaAla 56  
 DB 328 ATGGCCATATACGCCTAC-----TATTAACCGGGATTGGTGTGCTGTG 372  
 QY 57 LeuIlePheGlyTyTrileGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76  
 DB 373 CTATAGATTGGCTTACATCCAGGTTTCACTTGGTGTGCTGGCAGCTGGACAGACAGATACAC 432  
 QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
 DB 433 AAGATTAGGCAAGAGTTTTCATATGCTATAAATCAGGAGATGGCTGGTTTGTATGTG 492  
 QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr--AspIleAspLysIleSerAspGly 115  
 DB 493 CATGATGTTGGGAGCTCAACACCGCTGCACAGATGATGCTCCAAATTAATAGCGGA 552  
 QY 116 IleGlyAspLysIleAlaLeuLeuGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
 DB 553 ATTTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACACCAATTTTATAGCCGTTTATC 612  
 QY 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
 DB 613 ATAGGATTTTATAGTGGTTGGAAAGTCAACCCCTGTCATTTTGGCTGTCAGCCCTCTTATT 672  
 QY 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
 DB 673 GGAATTGTCATCTCTTGTGGGCAAGAGTATTGACTTCAATTACTAATAGGAACCTCCAG 732  
 QY 176 AlaTySerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195  
 DB 733 GCTTATGCAAAACCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792  
 QY 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 215  
 DB 793 GCCTTTGGAGGACACAGAAAGAACTTGA----- 822

QY 216 AlaTrpPheTyPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyThr 235  
 DB 823 -----AGGTACAAT 831  
 QY 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255  
 DB 832 AAAAATTTAGAAAGAGCTAAAAATGTTGCATAAAGAAATATACACGCCATCATTCG 891  
 QY 256 LeuGlyAlaValTyPheMetAsnGlyThrTyGlyLeuAlaPheTrpTyGlyThr 275  
 DB 892 ATAGGATTCCTCTCAAAATGAA-----TATTCTATTGGAGAAGTGTCTACTGTCTCTTC 951  
 QY 276 SerLeuIleLeuAsnGlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePhe 295  
 DB 952 TCTTTGGTCTCTCAAAATGAA-----TATTCTATTGGAGAAGTGTCTACTGTCTCTTC 1005  
 QY 296 SerValIleHisSerSerTyCysIleGlyAlaAlaValProHisPheGluThrPheAla 315  
 DB 1006 TCTATTTTGTGGGACTTTTATGATTGGACACTTGGCCCCCAACATAGAAGCTTTTGA 1065  
 QY 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335  
 DB 1066 AAGGCACGAGGGGCGACCTTTGAAATCTTCAAGATAATTGATAACGAGCCAGCATTTGAC 1125  
 QY 336 AsnPheSerThrAlaGlyTyTrpLysProGluSerIleGlyThrValGluPheLysAsn 355  
 DB 1126 AGCTTCTCAACAAAGGGCTACAAACCCAGACAGTATATATGGAAACTTAGAGTTTAAAT 1185  
 QY 356 ValSerPheAsnTyProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375  
 DB 1186 GTTCACTTCAATACCCATCGAAGAGGAGTTCAGATCTTGAAGGGCTCAATCTGAAG 1245  
 QY 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395  
 DB 1246 GTGAGAGCGGACAGACGGTGGCTTGGTTGGCAACAGTGGTGTGGAAGAACCAACT 1305  
 QY 396 ValGlnLeuLeuGlnArgLeuTyTrpAspProAspAspGlyPheIleMetValAspGluAsn 415  
 DB 1306 GTCCAGCTCATGACAGAGGCTTACGACCCCTGGAGGGCGTGTGATATCGAGGACAA 1365  
 QY 416 AspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyValValSerGlnGlu 435  
 DB 1366 GACATCAGAACCAATCAATGTAGGTATCTGAGGGAGATCAITTTGGTGTGATCAGGAA 1425  
 QY 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyGlyArgAspValThr 455  
 DB 1426 CCTGTCTGTTTCCACCCAGCATGCCGAGAACATTCGGCTATGGCCGAGAAAGATGTACC 1485  
 QY 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyAspPheIleMetGluPhe 475  
 DB 1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTTCATGAAACTG 1545  
 QY 476 ProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLys 495  
 DB 1546 CCCCACCAATTTGACACCCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGACAGAAA 1605  
 QY 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515  
 DB 1606 CAGAGAAATCGCATTTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTGGACGAG 1665  
 QY 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAla 535  
 DB 1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGTGTGTGTCAGGCCCTGATAGGCT 1725  
 QY 536 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 555  
 DB 1726 AGAAGAGCGGAGCACCACCATTTGTGATGATCATCGCTTGTCTACAGTTCGTATGTCTGAC 1785  
 QY 556 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 575  
 DB 1786 GTCATTTGCTGGTTTGTGATGGTGTGTGATTTGTGGACAGGAAGAAATCATGATGAGCTCATG 1845



;  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 7  
; LENGTH: 4788  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-866A-7

## Alignment Scores:

Pred. No.: 4,77e-311 Length: 4788  
Score: 3499.00 Matches: 668  
Percent Similarity: 73.13% Conservative: 252  
Best Local Similarity: 53.10% Mismatches: 264  
Query Match: 56.81% Indels: 74  
DB: 10 Gaps: 9

US-09-873-409-5 (1-1222) x US-09-866-866A-7 (1-4788)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
Db 148 ATGCTGGTGGAACTCTGGCTGTATATCCATGGAGTGGCGCTCCCACTTATGATGCTG 207
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
Db 208 ATCTTTGGTGACATGACAGATAGCTTTGGCAAGTGTAGGAAACGCTCTCTAAACACAGTACT 267
Qy 39 -----Thr 39
Db 268 AATPATGAGTGAGCGCCGATAAAGAGGCATGTTTGCCTCAAACTGGAGGAAGAAATGACCAAG 327
Qy 40 TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaIlePhe 59
Db 328 TAGCCCTAC-----TATTACCCGGGATGGTCTGGTGTGCTCATAGTT 372
Qy 60 GlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 373 GCCTACATCCAGGTTTCATTTTGGTGGCGAGTGGAGAGAGATGAGTGGTGGTGGTGGTGGTGG 432
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 433 CAGAAGTTTTTTTCATGATAATGAATCAGGAGATAGGCTGGTGGTGGTGGTGGTGGTGGTGGT 492
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 493 GGGGAGCTCAACCCGGCTCAGATGATGTTTCCAAATTAATGAAGGAATTTGGTGAC 552
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 553 AAAATCGGAATGTTCTCCAGGCAATGGCAACATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 612
Qy 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 613 ACCCGTGGGTGGAGCTAACCCCTGTGATTTTGGCCATCAGCCCTGTCTTGGCATGTGCA 672
Qy 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 673 GCTGGTATTGGGCAAGATATTGCTTCATTTACTGATAAGGAATCCATGCTTATGCA 732
Qy 179 LysAlaGlyValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 733 AAGCTGGAGCAGTCTCTGAGAGAGCTTAGCAGCCATCAGAACTGATGCTGATGGCTTGA 792
Qy 199 AlaGlnGlyGluGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 793 GGACAAAGAGGAAGTCTGAA-----813
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db 814 -----AGGTACAATAACAACTTG 831
```

```
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 832 GAAGAAAGCTAAAAGGCTGGGATAAAGAAAGCTATCACGCCCAACATCTCCATGGGTGCA 891
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 892 GCTTTTCTCTTATCATGCATCATATGCTCTGGCAATCTTGGTATGGGACTTCTCTTGGTC 951
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 952 ATCTCCAAAGAA-----TACTCTATTGGCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1005
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1006 ATTGGAGCATTCAGTGTGGCAGGCAATCTCCAAATATTGAAGCCTTCGCCAATGACACGA 1065
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1066 GGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCAGTATAGACACTTCTCA 1125
Qy 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
Db 1126 AAGAGTGGGCACAAACACACACATACAAAGNAATCTCGAATTTAAGAAATATTCACTTC 1185
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1186 AGTTACCCATCTCGAAAAGAAGTTCAAGTCTTGAAGGGCTCAATCTGAAGGTGAAGAGC 1245
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1246 GGACAGACGGTGGCTGGTGGTGGCAACAGTGGCTGTGGAAAGACACAACTGTCTCAGCTG 1305
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1306 ATGCAAGGCTCTACAGCCCTTAGATGGCATGGTCAGTATGACGACAGGACATCAGA 1365
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1366 ACCATCAATGTGAGGTATCTGAGGGAGATCATTTGGTGTGGTGGTGGTGGTGGTGGTGGT 1425
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db 1426 TTTGCCACCAACCATGCCGAGAACATTCCTATGGCCGAGAGATGTCCACATGGATGAG 1485
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1486 ATGAGAAAGCTGTCAAGGAAGCCATGCTATGATCTTCATGAAGACTGCCCCACCAA 1545
Qy 479 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1546 TTTGACACCCCTGGTGGTGAGAGAGGGGGCGCAGTGGTGGGGGACAGAAACAGAGAATC 1605
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1606 GCATTTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTGGACGAGGCCACTCA 1665
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGly 538
Db 1666 GCCTTGGATACAGAAAGTGAAGCTGTGGTTCAGGGCCGCACTGGATGAAGCTAGAGAGGC 1725
Qy 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db 1726 CGGACCAACCATTTGTAGTACGTCTCGTGTCTACCGTTCTGAATGCTACGCTCATTTGCT 1785
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db 1786 GGTTTTGTATGGTGTCTATCTGGAGCAAGGAATCATGATGAGTCTCATGAGAGAAAG 1845
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGln-----588
Db 1846 GGCATTACTTCAAACTTGTCTATGACACAGCAGCAGGAAATGAAATTGAATTAGGAAT 1905
Qy 589 AspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThr 608
```

```
Db 1906 GAAGCTTGTAAATCTAAGGATGAATTTGATATTTAGACATGCTCTTCAAAAGAGATTGAGGA 1965
Qy 609 AnpSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624
Db 1966 TCCAGTCTAATAAGAAGA-----AGATCAACTCGCNAAGACATCTGTGGACACACATGAC 2019
Qy 625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db 2020 CAAGACAGGAAGCTTAGTACAAAGAGGCCCTCGATGAAGATGATACCTCCAGCTTCTTT 2079
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db 2080 TGCCGATCTGAAGTTGAATCAACTGAATGCCCTATTCTTGTGGTGGTATATTCTGT 2139
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleLeuThr 683
Db 2140 GCCATAATAATGGAGCTTACAGCCAGCAATTTCTCGTATATATTTCAAAAGTTGTAGGG 2199
Qy 684 MetPheGlyAsn---AnpAspLysThrThrLeuLysHisAspAlaGluIleLysSerMet 702
Db 2200 GTTTTACAAATGGTGGCCCTGAAACCCAGCGGAGAACCAACTTGTGTTCTCTTG 2259
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db 2260 TTGTTTCTGATCTCTGGGATCATTTCTTCTTCAATACATTTTTTCTCAGGGCTTCACATTT 2319
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db 2320 GGCAAGCTGGAGATCTTCCACAGCGACTCCGATACATGTTTCAATTCATGCTG 2379
Qy 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db 2380 AGACAGGATGTGAGCTGTTTGTATGACCTTAAACACACCCAGGACCTGACCACAGG 2439
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db 2440 CTCGCCAACGATGCTGCTCAAGTGAAGGGGCTACAGGGTCTAGGGTGTGCTGTGATTTTC 2499
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 802
Db 2500 CAGAACATACCAATCTTGGGACAGGATCATCATATCCCTATCTATGCTGGCAACTA 2559
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db 2560 ACATTTTACTCTTAGCAATTTGACCCATCATCTGCGATAGCTGGAGTGGTGAATGAAA 2619
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAla 842
Db 2620 ATGTTGTCGGACAAGCACTCAAAAGATAAAGAGCACTAGAAAGTTCTGGAAAAGATTGCT 2679
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db 2680 ACGGAACAATGAAATTTCCGCACCTGTTGTCTCTTTGACTCGGGAGAGAGTTTGAA 2739
Qy 863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db 2740 ACATGTATGCCAGACTTGCAGATACCATACAGAAATCGGATGAAGAAGCACAGGTG 2799
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db 2800 TTTGGGATCAGTTCTCTCTCACCCAGGCCATGATGATTTTCTTATGCTGTGTTTTC 2859
Qy 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2860 CGGTTCGGTCCCTACTTGGTGACACAACTCATGACTTTTGAATGTTCTGTGTAGTA 2919
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 2920 TTCTCAGCTATTGCTTGTGGTGCATGGCAGTGGGCGAGTTCATTCGCTCCCTGAC 2979
Qy 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsn 962
```

```
Db 2980 TATGCGAAGAACACAGTGTGTCAGATCCCATCATCATCAGGATCATTTGAGAAAAACCCCGAG 3039
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db 3040 ATTGACAGCTACAGCAGCGAAGGCTTAAAGCCGAATATGTGTGAAGAAATGTGCAATTT 3099
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3100 AGTGGAGTGTGTTCAACTATATCCACCCAGCATCCAGTGTCTTCCGGGGCTGAGC 3159
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerGlyCysGlyLysSer 1022
Db 3160 CTTGAGTGAAGAAGGCCAGACGCTGGCCCTGTGGGAGCAGTGTGCTGGGGAAGAGC 3219
Qy 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db 3220 ACAGTGTCTCAGCTGCTCGAGCGCTTCTACGACCCCATCGCTGGATCAGTGTCTTAGAT 3279
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db 3280 GGCAAAAGAAATAAAGCACTGAATGTCCAGTGGCTCCGAGCACAGCTGGGCATTTGTGTC 3339
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3340 CAAGACCCCATCTCTTGTACTGCGACATCGCAGAGAAATTCCTCAGGAGCAACAGC 3399
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3400 CGGTCGTCTTATGAGGAGATTGTGAGGGCAGCCCAAGGAGGCCAATCCACCACTTC 3459
Qy 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3460 ATCGACTCGCTACCTCATTAATACAAACACAGTAGGAGACAAAGGCATTCAGCTGTGTC 3519
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeu 1142
Db 3520 GGTGGGCAAGCAGCGCATCGCCATCGCAGCGCCCTCGTACAGCAGCTCATATTITA 3579
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAla 1162
Db 3580 CTTCTGACGAGCAACATCAGCTCTGGATACAGAAAGTGAAAGGTTGTCCAGGAAGCG 3639
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db 3640 CTGGAACAAAGCCAGGGAAGCGCACCTGCTCATTTGTATCGCTCACCGCTGTCCACCATC 3699
Qy 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db 3700 CAGACCGGACTTGTATCTGTTGATTTCAGACGGCAAGGTCAAGGAGCAGCGACCCAC 3759
Qy 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3760 CAGCAGCTCTGCGCAGAAAGGCGCATCTACTTCTCAATGTCAGTGTGAGGCT 3813

RESULT 15
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
```

; SEQ ID NO 2299  
; LENGTH: 3924  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234  
US-09-880-107-2299

Alignment Scores:  
Pred. No.: 1,596-309 Length: 3924  
Score: 3481.00 Matches: 676  
Percent Similarity: 72.82% Conservative: 235  
Best Local Similarity: 54.04% Mismatches: 276  
Query Match: 56.52% Indels: 64  
DB: 10 Gaps: 9

US-09-873-409-5 (1-1222) x US-09-880-107-2299 (1-3924)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 201 ATGTCGCTGGTACCATCATGCGCATGCTCAGGATCAGGTCTCCCTCATGATGATA 260  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 261 GTATTGGAGAGATGACTGACAAATTTGTTGAT-----ACTGACGGAACCTTC 308  
Qy 41 SerPhe-----PheArgLeuThrLeu----- 47  
Db 309 TCTTTCCAGTGAACCTTTTCCTTGTCTGCTTAATCCAGGCNAATTTCTGGAGAGAA 368  
Qy 48 -----TyrTyrValGlyIleGlyValAlaLeuLeuPheGlyTyr 61  
Db 369 ATGACTAGATATGCATATTAATCTAGGATGGGTGCTGGAGTTCCTGTGCTGCTAT 428  
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleArgGlyGln 81  
Db 429 ATACAAGATTTTCATTTTGGACTTTGGCAGTGGTGCAGACATCAGGAAATTAGGCAGA 488  
Qy 82 PhePheHieSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 489 TTTTTCATGCTATTCTACGACAGGAATAGGATGGTTTGACATCAATGACACCCTGAA 548  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 549 CTCATACCGCGCTAACAGATGACATCTCCAAATCAGTGAAGAAATTTGTCGACAGGTT 608  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 609 GGAATGTTCTTTCAAGCAGTAGCCACGCTTTTTCGACAGGATTCATAGTGGGATTCATCAGA 668  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 669 GGATGAAGCTCACCTTGTGATAATGGCCATCAGCCCTATTCTAGACTCTCTGCGAGCC 728  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 729 GTTTGGCAAGATACCTCTCGCATTTAGTGACAAAGAACTAGCTGCTTATGCAAAAGCA 788  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 789 GGGCGCGTGGCAGAGAGGCTCTGGGGGCCATCAGGACTGTGTAGCTTTTCGGGGGCCAG 848  
Qy 201 GluLysGluLeuGlnArgSerPheLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 849 AACAAAGAGCTGGAA----- 863  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 864 -----AGGTATCAGAAACATTTTAGAAAT 887  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 888 GCCAAAGAGATTGGAATTTAAAGAGCTATTTCAGCAAAACATTTCCATGCTGATTTGCTTC 947

Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsn 280  
Db 948 CTGTTAATATATATGATCATATGCACTGGCTTCTGGTATGGATCCACTCTAGTATCA 1007  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1008 AAAGAA-----TATACTATTGGAAATGCAATGACAGTCTTTTTCATCTAATGGA 1061  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1062 GCTTTTCAGTGTGGCAGGCTGCCCATGTATTGATGCTTTTGCCAAATGCAAGAGAGCA 1121  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1122 GCATATGTGATCTTTGATATATTGATAATAATCTCTAAATTCACAGTCTTTTTCAGAGAGA 1181  
Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLeuPheLysAsnValSerPheAsnTyr 360  
Db 1182 GGACAAACCCAGCAGCATCAAGGAAATTTGGAGTTCAATGATGTTTCACCTTTCTTAC 1241  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1242 CTTTCTCGACTAAGCTCAAGATCTTGAAGGCTCACTGAAGGTCAGAGTGGGCAG 1301  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400  
Db 1302 ACGTGGCCCTGTTGGAAGTAGTGGCTGTGGAAAGAGCACAAACGCTCCAGCTGATACAG 1361  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1362 AGGCTCTATGACCTGATGAGGCAATTAACATGATGGCAGGATATTAGGAACCTTT 1421  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1422 AATGTAACATATCTGAGGAATCATTTGTTGGTGGTGCAGGCGCGTGTGTTTCC 1481  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1482 ACCCAATTTGCTGAAATATTTTATGGCCGTGGAATGTAAACCATGATGAGATAAAG 1541  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1542 AAAGCTGTCAAAAGAGCCCAACCCCTATGATGTTTATCATGAAATTAACACAGAAATTTGAC 1601  
Qy 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1602 ACCCTGTTGGAGAGAGAGGCGCCCGCTGAGTGGTGGCAGAGCAGAGGATCGCAT 1661  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1662 GCACGTGCCCTGTTCCGCAACCCCAAGATCTTCTGCTGATGAGGCCACGTCAGCAT 1721  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1722 GACACAAAGTGAAGCTAGGTACAGGAGCTCTGGATGAAGCCAGAGAGCCGAGCC 1781  
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1782 ACCATTGTATGACACACCGACTGCTACGGTCCGGAATGACAGATGTCATCGCTGGGTT 1841  
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1842 GAGGATGGAGTAATTTGTGGCAAGGAGGCCACAGCGAACTGATGAAGAGGAGGGGTG 1901  
Qy 581 TyrTyrSerLeuVal-----MetSerGlnAspIleLys 591  
Db 1902 TACTTCAAACTTGTCAACATCGACATCAGGAAGCCAGATCCAGTCAGAGAAATTTGAA 1961  
Qy 592 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 611  
Db 1962 CTAAATGATGAAGAGCTGCCACTAGAAATGGCCCCCAATGGCTGGAAATCTCGCCTATT 2021

```
QY 612 ProLeuHisSerVallySerIleLys---SerAspPheIleAspLysAla---GluGlu 629
Db 2022 AGGCATTCTACTCAGAAAACCTTTAAATAATTCACAAATGTGTGAGAAAGCCCTTGATGTG 2081
QY 630 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 649
Db 2082 GAAACCGATGGACTTGAAGCAATGTGCCACCAAGTGTCTTCTGAAAGGTCCTGAAACTG 2141
QY 650 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 669
Db 2142 AATAAACAAGATGGCCCTACTTGTGTGGGAACAGTATGTGCCATGTGCCATGGGGG 2201
QY 670 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAsp 699
Db 2202 CTTTCAGCCGGCATTTTCAGTCATATCTCAGAGATCATAGCATTTTGGACCCAGCGCAT 2261
QY 690 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 709
Db 2262 GATGCGAGTGAAGCAGCAGAAAGTGCACATATCTCTTTGATTTTCTTATTTCTGGGAATT 2321
QY 710 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 729
Db 2322 ATTTCITTTTACTTTCCTTCAGGTTTCACGGTTTCGGGAAGCTGGCAGATCCCTC 2381
QY 730 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 749
Db 2382 ACCAGAAGACTGGCGTCAATGGCTTTTAAAGCAATGTCTAAGACAGGACATGAGCTGGTTT 2441
QY 750 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 769
Db 2442 GATGACCATAAACAGTACTGTGTGCACCTTCTACAGACTTGCACAGATGCTGCCCAA 2501
QY 770 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 789
Db 2502 GTCCAAGGCCACAGCAGACAGGTTGGCTTTAATTCACAGATATAGTAACTTGGGA 2561
QY 790 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 809
Db 2562 ACTGGTATTATCATCATTTATCTACGGTTGGCAGTTAAACCTTATTTGCTATTAGCAGTT 2621
QY 810 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 829
Db 2622 GTTCCCAATTATGCTGTGTGAGAAATGTTGAAATGAAATGTTGGTGGAAATGCCAAA 2681
QY 830 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 849
Db 2682 AGAGATAAAGAAAGACTGGAGCTGTGGAAGATGCAACAGAGGCAATAGAAATATT 2741
QY 850 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 869
Db 2742 AGGACAGTTGTGCTTTGACCCAGGAAGAAAATTTGAATCAATGTATGTTGAAAAATTG 2801
QY 870 GlnThrGlnHisArgAsnThrSerLysLeuAlaGlnIleIleGlySerCysTyrAlaPhe 889
Db 2802 TATGGACCTTACAGGAATCTGTGCAAGAGCAGACATCTATGGAAATACTTTTAGTATC 2861
QY 890 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 909
Db 2862 TCACAGCATTTATGATATTTTCTATGCGGTGTTTTCGATTTGGTGATATCTCATTT 2921
QY 910 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 929
Db 2922 GTGAATGGACATATGCGCTTCAGAGATGTTATTTCTGTTGTTTCTGCAATTTGATTTGGT 2981
QY 930 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 949
Db 2982 GCAGTGGCTCTAGGACATGCCAGTTCATTGCTCCAGACTATGCTAAGCTAAGCTAAGCTGCT 3041
QY 950 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 969
Db 3042 GCAGCCCATTTATTCATGCTGTTTGAAGACCAACCTCTGATTGCACAGCTACAGTGAAGAG 3101
QY 970 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 989
```

```
Db 3102 GGCTGAAGCCCTGATAAATTTGAAGGAATAATAACATTTAATGAAGTCTGTGTCAACTAT 3161
QY 990 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 1009
Db 3162 CCCACCCGAGCAACCGTCCAGTCTTTCAGGGCTGAGCTTGAGGTGAAGAAGCCAG 3221
QY 1010 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGln 1029
Db 3222 ACACTAGCCCTGTGGCCAGCAGTGGCTGTGGGAAGACACGCTGTCTCCAGCTCTCTGGAG 3281
QY 1030 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1049
Db 3282 CGGTCTTACGACCCCTTGGCGGGACAGTCTCTCGATGGTCAAGAAAGCAAGAAATC 3341
QY 1050 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1069
Db 3342 NATGTCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCTATCTTATTGAC 3401
QY 1070 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1089
Db 3402 TGCAGCATTTGCCAGAAATATTCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGA 3461
QY 1090 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluLeuProGluLys 1109
Db 3462 ATTGTGAGTGCAGCCCAAGCTGCCAACATACATCTCTTTCATCGAGACGTTACCCCAAA 3521
QY 1110 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1129
Db 3522 TATGAACAAGAGTGGAGATTAAGGGAGCTCAGCTCTCAGGAGTCAAAAACAGAGGATT 3581
QY 1130 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1149
Db 3582 GCTATTGCCGAGCCCTCATCAGACAACCTCAATCTCTCTGTTGGATGAAGCTACATCA 3641
QY 1150 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1169
Db 3642 GCTCTGGATACTGAAAGTGAAGGTTGTCCAAGAACCCCTGGACAAAGCCAGAGAGGC 3701
QY 1170 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1189
Db 3702 CGCACCTGTATTGTGATTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTG 3761
QY 1190 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1209
Db 3762 GTGTTTCAGAAATGGGAGAGTCAAGGAGCAGTGGCAGCATCAGCAGCTGTCTGGCAGAAA 3821
QY 1210 AspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3822 GGCATCTATTTTTCATATGTCAGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 16:09:09

Job time : 711.798 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 178.956 Seconds  
(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-5  
Perfect score: 6159  
Sequence: 1 MILGILASLVNGACPLMLP.....QELLENRDYIFKLVAQSVQ 1222

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODE=frame+p2n.model -DEV=xlh  
-Q/cgn2\_1/USFTO\_spool/US09873409/runat\_27032003\_115422\_19259/app\_query.fasta\_1.7544  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CNC 1.1 578 @runat\_27032003\_115422\_19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA :  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCUTS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3521.5	57.2	4646	1 US-08-181-471-2	Sequence 2, Appli
2	3521.5	57.2	6505	1 US-08-793-610-5	Sequence 5, Appli
3	3521.5	57.2	9318	2 US-08-793-610-6	Sequence 6, Appli
4	3520	57.2	4264	2 US-08-784-649A-1	Sequence 1, Appli
5	3520	57.2	4264	2 US-08-784-649A-5	Sequence 5, Appli
6	3519.5	57.1	4669	6 5206352-3	Patent No. 5206352
7	3503.5	56.9	4669	2 US-08-752-447-1	Sequence 1, Appli
8	3503.5	56.9	4669	4 US-09-316-167-1	Sequence 1, Appli
9	3494	56.7	4233	3 US-09-120-513-1	Sequence 1, Appli
10	3494	56.7	4233	4 US-09-450-105-1	Sequence 1, Appli
11	3473.5	56.4	4669	2 US-08-583-276-18	Sequence 18, Appli
12	2280	37.0	2726	1 US-08-461-823-1	Sequence 1, Appli

13	2259	36.7	4002	2 US-08-996-545-1	Sequence 1, Appli
14	2259	36.7	4002	2 US-08-996-545-3	Sequence 3, Appli
15	2259	36.7	4002	4 US-09-328-320-1	Sequence 1, Appli
16	2259	36.7	4002	4 US-09-328-320-3	Sequence 3, Appli
17	2250	36.5	4047	2 US-08-612-734B-1	Sequence 1, Appli
18	2187.5	35.5	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	2143.5	34.8	3924	1 US-08-995-246C-1	Sequence 1, Appli
20	2100.5	34.1	4224	1 US-08-612-521-1	Sequence 1, Appli
21	1722	28.0	6143	1 US-08-612-521-3	Sequence 3, Appli
22	1680.5	27.3	3924	2 US-08-996-644-3	Sequence 3, Appli
23	1680.5	27.3	3924	3 US-09-352-552-3	Sequence 3, Appli
24	1680.5	27.3	3927	2 US-08-996-644-1	Sequence 1, Appli
25	1680.5	27.3	3927	3 US-09-352-552-1	Sequence 1, Appli
26	1646.5	26.7	3909	1 US-08-232-537-1	Sequence 1, Appli
27	1636	26.6	3792	4 US-09-351-224E-10	Sequence 10, Appli
28	1501.5	24.4	3999	4 US-09-351-224E-9	Sequence 9, Appli
29	1106.5	18.0	13188	4 US-08-961-527-70	Sequence 70, Appli
30	1047.5	17.0	7760	4 US-08-961-527-63	Sequence 63, Appli
31	989	16.1	441529	4 US-09-103-840A-1	Sequence 1, Appli
32	987	16.0	4403765	4 US-09-103-840A-2	Sequence 2, Appli
33	940	15.3	6492	4 US-08-961-527-188	Sequence 188, Appli
34	938	15.2	7186	4 US-08-961-527-39	Sequence 39, Appli
35	738	12.0	2376	1 US-08-394-880B-1	Sequence 1, Appli
36	708.5	11.5	1749	4 US-09-134-001C-1893	Sequence 1893, Ap
37	695	11.3	5889	1 US-08-463-092B-5	Sequence 5, Appli
38	695	11.3	5889	2 US-08-462-109A-5	Sequence 5, Appli
39	695	11.3	5889	2 US-08-460-907B-5	Sequence 5, Appli
40	695	11.3	5889	3 US-08-463-179A-5	Sequence 5, Appli
41	695	11.3	5889	3 US-08-461-384B-5	Sequence 5, Appli
42	687	11.2	5011	1 US-08-463-092B-3	Sequence 3, Appli
43	687	11.2	5011	2 US-08-462-109A-3	Sequence 3, Appli
44	687	11.2	5011	2 US-08-460-907B-3	Sequence 3, Appli
45	687	11.2	5011	3 US-08-463-179A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-08-181-471-2  
; Sequence 2, Application US/08181471  
; Patent No. 5641508  
; GENERAL INFORMATION:  
; APPLICANT: Li, Lingna  
; APPLICANT: Lishko, Valeryi K.  
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Thomas Fitting  
; STREET: 12526 High Bluff Drive, Suite 300  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92130  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/181,471  
; FILING DATE: 13-JAN-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/041,553  
; FILING DATE: 02-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: ANT0029P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-792-3680  
; TELEFAX: 619-792-8477



## INFORMATION FOR SEQ ID NO: 2:

\* SEQUENCE CHARACTERISTICS:  
LENGTH: 4646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4267  
US-08-181-471-2

## Alignment Scores:

Pred. No.: 0 Length: 4646  
Score: 3521.50 Matches: 681  
Percent Similarity: 73.92% Conservative: 246  
Best Local Similarity: 54.31% Mismatches: 264  
Query Match: 57.18% Indels: 63  
DB: 1 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-181-471-2 (1-4646)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 575 ATGGTGGTGGAACTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 634  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTC 694  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 695 AACATCACTAATAGAGTATATCAATACAGGGTCTTCATGAATCTGGAGGAGAC 754  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaLeuIlePheGlyTyr 61  
DB 755 ATGACAGGATGCTATATATACAGTGGAAATGGTCTGGGGTGGTGGTCTGCTTAC 814  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysAlaArgLysGln 81  
DB 815 ATTCAGGTTTCATTTGGTCTGGCAGCTCGGAACACAAATACACAAAATTAGAAACAG 874  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 875 TTTTTCATGCTATAATGCGACAGGATAGGCTGGTGTGATGTCACGATGTTGGGAG 934  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 935 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATGTGACAAAT 994  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 995 GGAATGTTCTTCAGTCAATGGCAACATTTTTCAGTGGGTTTATAGTAGGATTACACGT 1054  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 1055 GGTGGAAAGCTAACCCCTGTGATTTTGGCCATCAGTCTGCTTCTGGAGCTGTGCTGCT 1114  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 1115 GTCTGGCAAGATACATCTTCAATTTTACTGTATTAAGAACTCTTAGGCTATGCAAGCT 1174  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 1175 GGACAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATGATTTGGAGGACAA 1234  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 1235 AAGAAAGAACTTGA----- 1249  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240

DB 1250 -----AGGTACAAACAAAAATTTAGAAAGAA 1273  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGGTGCTGCTTTC 1333  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
DB 1334 CTGCTGATCATGCTATGCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCA 1393  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 1394 GGGGAA-----TATTCTATTGGACAAAGTACTGCTGTTCTCTCTGTTTCTGTTTAAATGGG 1447  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 1448 GCTTTTGTAGTTGGACAGGCATCTCCAGCATTTGAAGCAATTTGCAAAATCAGAGAGGCA 1507  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 1508 GCTTATGAATCTTCAAGATATTTGATATAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1567  
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 1568 GGGCACAAACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1628 CCATCTCGAAAAAGAAATTAAGATCTTGAAGGCGCTGAACCTGAAGGTGCGAGTGGGCA 1687  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1688 ACGTGGCCCTGGTGGAAACAGTGGCTGGGAAGAGACCAACAGTCCAGCTGATGCGAG 1747  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
DB 1748 AGGCTCTATGACCCACAGAGGGGATGGTTCAGTGTGTGATGGACAGGATATTAGGACCAT 1807  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1808 AATGTAAGGTTTCTACGGGAAATCATTTGTTGGTGGTACAGGAACTGTATTGTTTGGC 1867  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 1868 ACCACATAGCTGAAACCAATTCGCTATGCGCGTGAATAATGTCAACATGATGATGAG 1927  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1928 AAGAGTGTCAAGGAAGCAATGCTATGACTTTATCATGAAACTGCTCATAAATTTGAC 1987  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1988 ACCCTGTTTGGAGAGAGAGGGGCCAGTTGAGTGGTGGGCAAGAGCAGAGGATCGCCATT 2047  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 2048 GCACGTGCCCTGTTTCCCAACCCCAAGATCTCTCTGCTGGATGAGGCCAGCTCAGCTTG 2107  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
DB 2108 GACACAGAAAGCGAAGCAGCTGGTTTCAAGTGGCTCTCGATAAGGCCAGAAAGGTCCGACC 2167  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
DB 2168 ACCATTGTGATGCTCATCTGTTTGTCTACAGTTCGTAATGCTGACGCTATCGCTGTTTC 2227  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
DB 2228 GATGATGGAGTCAATGTGGAGAAAGGAATCATGTAATCATGTAATCATGAAAGAGAGGCAATT 2287  
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594



COUNTRY: U.S.A.  
 ZIP: 20005-5701  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/793,610  
 FILING DATE: 07-MAR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE P 44 31 973.8  
 FILING DATE: 08-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 195 03 952.1  
 FILING DATE: 07-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP95/03175  
 FILING DATE: 10-AUG-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bertram, Richard J.  
 REGISTRATION NUMBER: 39,105  
 REFERENCE/DOCKET NUMBER: P1614-7007  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 638-5000  
 TELEFAX: (202) 638-4810  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 6505 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA  
 US-08-793-610-5

## Alignment Scores:

Pred. No.: 0 Length: 6505  
 Score: 3521.50 Matches: 681  
 Percent Similarity: 73.92% Conservative: 246  
 Best Local Similarity: 54.31% Mismatches: 264  
 Query Match: 59.18% Indels: 63  
 DB: 2 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-793-610-5 (1-6505)

QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
DB	2387	GGAAATGCTTTTCAGTCAATGGCAACATTTTTCAGTGGGTTTATAGTAGGATTTACACGT	2446
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
DB	2447	GGTTGAAGCTAACCCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGTGT	2506
QY	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSerLysAla	180
DB	2507	GTCTGGGCAAGATACATCTCTTCTATCTGATAAAGAACTCTTAGCGTATGCAAAAGCT	2566
QY	181	GlyValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
DB	2567	GGAGCAGTAGCTGAAGAGCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA	2626
QY	201	GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe	220
DB	2627	AAGAAAGAACTTGAA-----	2641
QY	221	ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp	240
DB	2642	-----AGGTACAACAAATTTAGAGAA	2665
QY	241	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	260
DB	2666	GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGTGTGCTTTC	2725
QY	261	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTrpTyrGlyThrSerLeuIleLeuAsn	280
DB	2726	CTGCTCATCTATGCATCTTATGCTCTGGCTTCTGTTATGGGACCACTTGTGCTCTCA	2785
QY	281	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	300
DB	2786	GGGAA-----TATCTATTGGCAAGTACTCATCTATTCTTTCTGTTATTAAATGGG	2839
QY	301	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	320
DB	2840	GCTTTTAGTGTGGACAGCATCTCCAGCAATTTGAAGCAATTTGCAATTCGAGAGGACA	2899
QY	321	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	340
DB	2900	GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAATTTGACAGACTATTCCGAAGAT	2959
QY	341	GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr	360
DB	2960	GGGCAACAAACCATGATAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTAC	3019
QY	361	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	380
DB	3020	CCATCTCGAAAGAAAGTTAAGATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAG	3079
QY	381	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	400
DB	3080	ACGGTGGCTGTTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGATGTCAG	3139
QY	401	ArgLeuTyrAspProAspAspGlyPheIleMetValAspLysAsnAspIleArgAlaLeu	420
DB	3140	AGGCTCTAGACCCACAGAGGGGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3199
QY	421	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	440
DB	3200	AAATGTAAGGTTTCTACGGGAATCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3259
QY	441	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu	460
DB	3260	ACCACGATAGCTGAAACATTCGCTATGCGTGTGAAATGTCCACCATGGATGAGATTGAG	3319
QY	461	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	480
DB	3320	AAAGTGTCAAGGAAGCCCAATGCTTATGCTTATCATGAAATGCTCATAAATTTGAG	3379
QY	481	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	500

||||| 3380 ACCCTGTTGGAGAGAGGGGCCAGTTCAGTGGTGGGAGAGAGAGGATCGCCATT 3439  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleuAspGluAlaThrSerAlaLeu 520  
Db 3440 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTCTGCTGATGAGGCCAGCTCAGCCTTG 3499  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 3500 GACACAGAAAGCGAAGCAGTGGTTCAGTGGCTCTGGATAAGGCCAGAAAGTCGGACC 3559  
QY 541 ThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 3560 ACCATTGTGATAGTCATCGTTTGTCTACAGTTCTGTAATGCTCAGCTCGCTGTTTC 3619  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 3620 GATGATGGAGTCATTTGGTGGAAAGGAATCATGATGACTCATGAAGAGAAGGCATT 3679  
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
Db 3680 TACTTCAAACTTGTCAATGCACAGCAGCAGGAAATGCAAGTTCAATTAGAAAATGCAGCT 3739  
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 3740 GATGAATCCAAAGTGAATTCGATGCTTGGAAATGCTTCAAAATGATTCAAGATCCAGT 3799  
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
Db 3800 CTAAATAGAAAAAGATCAACTCTAGAGGTGCTCGTGATCAGCAAGCCCAAGACAGAAAG 3859  
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647  
Db 3860 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTGGAGGATTATG 3919  
QY 648 LysLeuAsnLysProGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
Db 3920 AAGCTAAATTTAATCAATGAGTCTGATTTTGTGGTGTATTTTGTGGTGTATTTTGTGGCATTATAAAT 3979  
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687  
Db 3980 GGAGGCTGCACACAGATTTGCAATATATTTTCAAGATATATAGGGTTTTACAGA 4039  
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706  
Db 4040 ATTGATGATCTCGAACAACACACAGAAATAGTAACTGTTTTCACATTATTTCTAGCC 4099  
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726  
Db 4100 CTTGGAAATTATTTCTTTTATTACATTTTCTTCCAGGGTTTCACATTTGCGAAAGCTGGA 4159  
QY 727 GluIleuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746  
Db 4160 GAGATCCTCAACAGGGCTCCCATATAGTTTTCGATCCATGCTCAGACAGGATGTG 4219  
QY 747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766  
Db 4220 AGTTGGTTTGATGACCCCTAAACACACACTGGAGCATTGACTACCGGCTCGCCAAATGAT 4279  
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786  
Db 4280 GCTGCTCAAGTTAAAGGGCTATAGGTTTCAGGCTTCTGCTGTAATTACCCAGAAATATAGCA 4339  
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806  
Db 4340 AATCTTTGGGACAGGAATAATATATCTCTCATCTATGCTTGGCAACTTAACACTGTTTACTC 4399  
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826  
Db 4400 TTAGCAATTTGATCCCATCTTGCATATAGCAGGAGTGTGTAATGAATGAATGTTGTCTGGA 4459  
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846  
|||||

Db 4460 CAAGCAGCTGAAGATAGAAGAACTAGAAAGTGTGGGAAGATCGCTACTAGCAATA 4519  
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866  
Db 4520 GAAACTTCCGAACCGTGTCTTTCTTACTCAGGAGCAGAACTTTGAACATATGTATGCT 4579  
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886  
Db 4580 CAGAGTTTGCAGGTACCATACAGAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 4639  
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906  
Db 4640 TTTTCTTCCACCAGCAATGATGATATTTTCTATGCTGGATGTTTCCGGTTTGAGCC 4699  
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926  
Db 4700 TACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGT 4759  
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946  
Db 4760 GTCTTTGGTGGCCTGCGCTGGGCAAGTCAGTTTCATTTTGTCTCTGACTATGCCAAGCC 4819  
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966  
Db 4820 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACACTAC 4879  
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986  
Db 4880 AGCAGGAGGCTTAATGCCGAACACATTCAGGAGGAAATGTCACATTTGGTGAAGTTGTA 4939  
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006  
Db 4940 TTCAACTATCCACCCGACCGGACATCCAGTCTTTCAGGAGCTGAGCCTGGAGGTGAAG 4999  
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026  
Db 5000 AGGGCCAGCAGCTGGCTCTGGTGGCAGCAGTGGCTGTGGGAAGACACAGTGTGTCAG 5059  
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046  
Db 5060 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAGTCTGCTTCATGGCAAGAAATA 5119  
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066  
Db 5120 AAGCGACTGAATGTTTCAGTGGCTCCGAGCACCTCGGCATCTGTGCCAGGAGCCATC 5179  
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086  
Db 5180 CTGTTTGACTGCAGCATGCTGAGAACTTGCCTATGGAGACACAGCCGGTGGTGTCA 5239  
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106  
Db 5240 CAGGAAGAGATCGTAGGCGAGCAGCAAGAGGCGCAACATATACCTTCATCGAGTCACTG 5299  
QY 1107 ProGluLysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126  
Db 5300 CCTAATAATATAGCACTAAAGTAGGAGACAAAGAACTCAGCTCTCTGTGTGCCAGAA 5359  
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146  
Db 5360 CAACGATTTGCCATAGCTCTGCTGCTTGTAGACCCCTCATATTTTGTCTTTGGATGAA 5419  
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166  
Db 5420 GCCACGTGAGCTCTGATACAGAAAGTGAAGGTTGTCCNAGAAGCCCTGGACAAAGCC 5479  
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186  
Db 5480 AGAAGAGCGCGCACCTGCTGATTTGTTGTTGCTCAGCCCTGTCCACCATCCAGATGCAGAC 5539  
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206  
Db 5540 TTAATAGTGTGTTTTCAGAAATGGCAGAGTCAAGGAGCATGGCCACGCATCAGCAGCTGCTG 5599

QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 5600 GCACAGAAAGGCATCTATTTTCAATGCTAGTCCAGGCT 5641

RESULT 3  
US-08-793-610-6  
Sequence 6, Application US/08793610  
Patent No. 5858744  
GENERAL INFORMATION:  
APPLICANT: BAUM, Christopher  
APPLICANT: STOCKING-HARBERS, Carol  
APPLICANT: OSTERTAG, Wolfram  
TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/793,610  
FILING DATE: 07-MAR-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 31 973.8  
FILING DATE: 08-SEP-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: DE 195 03 952.1  
FILING DATE: 07-FEB-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP95/03175  
FILING DATE: 10-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bertram, Richard J.  
REGISTRATION NUMBER: 39,105  
REFERENCE/DOCKET NUMBER: P1614-7007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)638-5000  
TELEFAX: (202)638-4810

INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA  
US-08-793-610-6

Alignment Scores:  
Pred. No.: 0 Length: 9318  
Score: 3521.50 Matches: 681  
Percent Similarity: 73.92% Conservative: 246  
Best Local Similarity: 54.31% Mismatches: 264  
Query Match: 57.18% Indels: 63  
DB: 2 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-793-610-6 (1-9318)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 1926 ATGGTGGTGGGAACCTTTGGTGCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 1985

QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 1986 GTGTTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAGATCTGATGCTCA 2045

QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 2046 AACATCACTAATAAGAGTGTATCATGATACAGGGTTCTTTCATGAATCTGGAGGAAGAC 2105  
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 2106 ATGACCAGGTATGCTTATTATACAGTGAATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2165  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaAaGlnThrLysArgIleAaGlyGln 81  
Db 2166 ATTCAGGTTTCATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2225  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 2226 TTTTTCATGCTATATATGACAGAGATAGGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2285  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 2286 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTTGGTGACAAAATT 2345  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 2346 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 2405  
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 2406 GGTGGAGAGCTAACCTTGTGATTTGGCCATCAGTCTCTGTCTTGGACTGTGAGTGTCT 2465  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 2466 GTCCTGGCAAGATACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2525  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 2526 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTTAGAACTTGTGATTCATTTGGAGGACAA 2585  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220  
Db 2586 AGAAGAGACTTGA----- 2600  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 2601 -----AGGTACAAACAAAAATTTAGAGAA 2624  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 2625 GCTAAAGAGAAATGGGATAAAGAAAGCTATTACAGCAATATTCTATAGGTGCTGCTTTC 2684  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 2685 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGGTATGGGACCACCTTGTCTCTCTCA 2744  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 2745 GGGGAA-----TATTCTATTGGCAAGTACTCAGTGTATTTCTTTCTGTTAATTTGGG 2798  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaGlyAla 320  
Db 2799 GCTTTTAGTGTGGACAGGCTCTCCAAAGCATTTGCAAAATGCAAGAGAGCA 2858  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 2859 GCTTATGAATCTTCAAGATATTGATTAATAGCCAAATTTGACAGCTATTGCAAGAGT 2918  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 2919 GGGCACAACCAAGATATATTAAAGGAAATTTGGAATTTGCAAAATGTTCACTTCACTTAC 2978  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 2979 CCATCTCGAAAAAGAGTAAAGATCTTTGAAGGGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3038





Db 5199 CAGGAAGAGATCGTGGGCGAGCAAGAGGCGCAACATACATACGCTTCTCATGAGTCACTG 5258  
QY 1107 ProGluysrYrAenThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLys 1126  
Db 5259 CCTAATAATATAGCTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGGTGCCAGAAA 5318  
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146  
Db 5319 CAACGCATTGCCATAGCTCGTCCCTGTTAGACAGCCTCATATTTGCTTTGGATGAA 5378  
QY 1147 AlaThrSerAlaLeuAspAenAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166  
Db 5379 GCCACGTCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAGAGGCCCTGGCAAGGCC 5438  
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAenAlaAsp 1186  
Db 5439 AGAAGAGGCCGACCTGTCATGTGATTGCTCCAGCCGCTGCCACCATCCAGAAATGCAGAC 5498  
QY 1187 LeuIleValValLeuHisAenGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeu 1206  
Db 5499 TTAATAGTGTGTTTCAGATGGCAGAGTCAAGAGCATGGCCAGCATCAGCAGCTGCTG 5558  
QY 1207 ArgAsnArgAspIleYrPheLysLeuValAsnAlaGlnSer 1220  
Db 5559 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 5600

## RESULT 4

US-08-784-649A-1  
; Sequence 1, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; APPLICANT: Chen, Gang  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,649A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: Reg.No. 5830697 36,677  
; REFERENCE/DOCKET NUMBER: 06037/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5070  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4264 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-784-649A-1  
Alignment Scores:  
Pred. No.: 0 Length: 4264  
Score: 3520.00 Matches: 682  
Percent Similarity: 74.00% Conservative: 246  
Best Local Similarity: 54.39% Mismatches: 262

Query Match: 57.15% Indels: 64  
DB: 2 Gaps: 11  
US-09-873-409-5 (1-1222) x US-08-784-649A-1 (1-4264)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20  
Db 292 ATGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACCTTCCTCTCATGATGCTG 351  
QY 21 ValLeuGlyGlyMetSerAsp-----AsnLeuIleSer 31  
Db 352 GTGTTTGGAGAATGACAGATATCTTTGCAATGCGAGAAATTTAGAGATCTGATGTCA 411  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 412 AACATCACTAATAGAGTGTATCATGATACAGGCTTCTTCATGAATCTGGAGGAAGAC 471  
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 472 ATGACCAGGTATGCCTATTATTACAGTGGAAATGGTCTGGGTGCTGGTGTGCTGTAC 531  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 532 ATTCAGGTTTCATTTTGGTGCCTGCGCAGCTGGAAGCAAAATACACAAATTTAGAAACAG 591  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 592 TTTTTCATGCTAATAGCAGGAGATAGGCTGGTTGATGTGCACCATGTGGGGAG 651  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAAATGAAGGAATTTGTCACAAAAT 711  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 712 GGA---ATGTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGAATTTACCGT 768  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 769 GGTGGAGACTAACCTTGATTTGGCCATCAGTCTCTGTTCTGGACTGTGACGTGCT 828  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 829 GTCTGGGCAAGATACTATCTTCAATTACTATAAAGAACTCTTAGCGTATGCAAAAGCT 888  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTTAGAACTGTGATTGCATTTGGAGGACAA 948  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 949 AGAAGAACTTGA----- 963  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 964 -----AGGTACAAACAAAAATTTAGAAAGAA 987  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 988 GCTAAAGAAATTTGGGATNAAGAAAGCTATTACAGCAATATTCTATAGTGTGCTTTC 1047  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
Db 1048 CTGCTCATCTATGCATCTTATGCTCTGGCTTCTGGTATGGGACCACCTTGGTCTCTCA 1107  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1108 GGGGAA-----TATTCATTGGCAAGTACTCAGCTGTATTCTTTCTGTTAATTAATGGG 1161  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1162 GCTTTTAGTGTGGACAGGCATCTCCAGCATTTGAAGCATTTTGAAGCATTTGAAGGAGCA 1221





Db 3382 CTCCTGAGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGTCATGGCAAGAAATA 3441  
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066  
Db 3442 AAGCGACTGAATGTTTCAGTGGCTCCGAGACACCTGGGATCTGTGTCAGGAGGCCATC 3501  
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086  
Db 3502 CTGTTTACGTGACGATGCTGAGAACATTTGCTATGGAGACAACAGCCGGTGGTGTCA 3561  
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106  
Db 3562 CAGGAAGAGATTGTGAGGCGAGCAAGAGGAGGCCAACATACATCGCTTCATCGAGTCACTG 3621  
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGluAlaGlnLeuSerGlyGlyGlnLys 1126  
Db 3622 CCTAATAATATAGCTAATAAGTAGGAGACAAAGGAACCTCAGCTCTCTGGTGGCCAGAA 3681  
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146  
Db 3682 CAACGCAATGCCATAGCTGCTGCGCTTGTAGACAGCCTCATATTTTGTCTTTGGATGAA 3741  
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166  
Db 3742 GCCACGTCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAGAGCCCTGGCAAGGCC 3801  
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186  
Db 3802 AGAAGAGCCGCACTGCTGATGTTGATTCACCGCTGTCACCATCCAGATCCAGATGGAGAC 3861  
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206  
Db 3862 TTAATAGTGTGTTTCAGATGTCAGAGTCAAGGAGTCAAGGAGTGGCAGCATCAGCAGCTGCTG 3921  
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 3922 GCACAGAAAGGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 3963

## RESULT 5

US-08-784-649A-5  
Sequence 5, Application US/08784649A  
Patent No. 5830697  
GENERAL INFORMATION:  
APPLICANT: Sikic, Branimir I  
APPLICANT: Chen, Gang  
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fish & Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,649A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: Reg.No. 5830697 36,677  
REFERENCE/DOCKET NUMBER: 06037/007001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 4264 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-784-649A-5  
Alignment Scores:  
Pred. No.: 0  
Score: 3520.00  
Length: 4264  
Matches: 682  
Percent Similarity: 74.00%  
Best Local Similarity: 54.39%  
Query Match: 57.15%  
Indels: 64  
Gaps: 11

US-09-873-409-5 (1-1222) x US-08-784-649A-5 (1-4264)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 292 ATGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACCTTCCTCTCATGATGCTG 351  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 352 GTGTTGGAGAAATGACAGATATCTTGCAAATGCGAGAAATTTAGAACATCTGATGTCA 411  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 412 AACATCACTAATAAGAGTGCATCAATGATACAGGGTCTCTCATGAATCTGGAGGAAGAC 471  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 472 ATGACCCGATATGCTTATTATACAGTGGAAATTTGGTGGGGTGTCTGCTGCTTAC 531  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 532 ATTCAGTTTCATTTTGGTGCTGCAGCTGGAGCAAAATACACAAATTTAGAAACAG 591  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 592 TTTTTCATGCTAATAATGCAGCAGGAGATAGCTGGTTGATGTGCACGATGTTGGGAG 651  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATAGAGGAATTTGGTCACAAAAT 711  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 768  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 769 GGTGGAAGCTAACCCCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTTCAGCTGCT 828  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 829 GTCTGGGCAAGATACATCTCTTCATTTACTGATAAAGAACTCTTACCGTATGCAAAAGCT 888  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 889 GGAGCAGTAGCTGAGAGGCTCTGGCAGCAATTTAGAACATTTGATTCATTTGGAGACAA 948  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 949 AAGAAAGAACTTGAA----- 963  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 964 -----AGGTACAAACAAAAATTTTAGAAGAA 987  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 988 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCCAATATTTCTATAGTGTCTGCTTTC 1047

QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuLeuLeuAsn 280  
Db 1048 CTGCTGATCATCTTATGCTTGTGGCTTCTGATGATGGACCACTTGTGCTCTCA 1107  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1108 GGGGAA-----TATTCTATTGGAAGAAGTCACTGTATTTCTTCTGTATTAAATGGG 1161  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1162 GCTTTTGTAGTGGACAGCATCTCCAAGCATTTGAAGCATTTGCCAAATGCAAGAGGAGCA 1221  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1222 GCTTATGAAATCTTCAAGATAATTGATAAAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1281  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1282 GGGCACAACCAAGATTAATTAAGGGAATTTGGAAATTCAGAAATGTTCACTTCAGTTAC 1341  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1342 CCATCTCGAAGAAAGATTGAATCTTGAAGGCTGAACCTGAAGTGCAGAGTGGGCAG 1401  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400  
Db 1402 ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGAGACACACAGTCCAGCTGATGCAG 1461  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1462 AGGCTCTATGACCCACAGAGGGAGTGGTCAGTGTGTGAGCAGGATATTAGGACCATA 1521  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1522 ATGTGAAGTTTCTACGGGAATCATTTGTGTGTGTGAGTGCAGGAACCTGTATTGTGGC 1581  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460  
Db 1582 ACCACGATAGCTGAAACATTCGCTATGCGGTGAAATGTCCACCATGATGAGATTGAG 1641  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1642 AAAGCTGTCAAGGAAGCCAAATGCTATGATCTTATCATGAACCTGCCTCATAAATTTGAC 1701  
QY 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1702 ACCCTGGTTGGAGAGAGAGGGGCCAGTGTGAGTGTGGGAGAGCAGAGGATCGCCATT 1761  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1762 GCAGTGGCCCTGTTGCGAACCCCAAGATCCTCTGCTGGATGAGGCCACGTCAGCCTTG 1821  
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1822 GACACAGAAGCGAAGCAGTGGTTCAGTGGCTCTGGATAAGGCCAGAAAGTCCGAGCC 1881  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1882 ACCATTGTGATGCTCATCTTTGCTACAGTTTCGTAATGCTCAGCTCAGCTCGCTGTTTC 1941  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1942 GATGATGGAGTCATTGTGGAGAAAGGAATCATGATGAATCATGAAGAGAAAGGCATT 2001  
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
Db 2002 TACTTCAAACTTGTCAATTCAGACAGCAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2061  
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 2062 GATGAATCCAAAAGTGAATGTGCTGGAAATGTCTCAATGATTCAGATCCAGT 2121  
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627

Db 2122 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCAAGCCCAAGACAGAAAG 2181  
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647  
Db 2182 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTTGGAGGATTATG 2241  
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
Db 2242 AAGCTAAATTTAACTGAATGGCTTATTGTTGTGTGTATTGTTGGCCATTATAAAT 2301  
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleMetPheGlyAsn 687  
Db 2302 GGAGGCTCTCAACCCAGCATTTGCAATAATTTTCAAGATATTAGGGTGTTTTACAAGA 2361  
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706  
Db 2362 ATTGATGATCTCTGAACAAACACGACGATAGTAACTGTTTTCTACTATTGTTCTAGCC 2421  
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726  
Db 2422 CTTGGAATTTATTTCTTTTATTACATTTTCTTCCAGGTTTCCACATTTGCGCAAGCTGA 2481  
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746  
Db 2482 GAGATCTCTCACAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACGAGTGTG 2541  
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766  
Db 2542 AGTGTGTTTGTAGACCTTAAACACCACTGGAGCATTGACTACAGGCTCGCCAATGAT 2601  
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786  
Db 2602 GCTGCTCAAGTTAAAGGGCTATAGTTCAGGCTTGTGTAATTAATGAAATGTTGCTGGA 2661  
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806  
Db 2662 AATCTTGGGACAGGAATAATTATCTTTCATCTATGTTGGCACTAAACACTGTTTACTC 2721  
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826  
Db 2722 TTAGCAATTTGCCATCATCTGCAATGACGAGGATGTTTGAATGAAATGTTGCTGGA 2781  
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846  
Db 2782 CAAGCACTGAAAGATAAGAAAGAACTAGAAAGTGTCTGGGAAGATCGCTACTGAAGCAATA 2841  
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866  
Db 2842 GAAAACTTCCGAACCGTGTGTTCTTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGCT 2901  
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886  
Db 2902 CAGAGTTTCAGATCCATPACAGAAACTCTTTGAGGAAGACACATCTTTGGAATTACA 2961  
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906  
Db 2962 TTTTCTTCCACCAGCAATGATGATTTTCTATGCTGGATGTTTCCGGTTGGAGCC 3021  
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926  
Db 3022 TACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTGTGTATTATTTTTCAGCTGT 3081  
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946  
Db 3082 GTCTTGGTGCCATGGCCGTGGGCAAGTCAGTTTCATTTGCTCTCTACTATGCCAAGCC 3141  
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966  
Db 3142 AAAATATCAGAGGCCCAATCATCATGATCATTTGAAATAAACCCCTTTGATTGACAGCTAC 3201  
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986

Db 3202 AGCAGGAGGCTTAATGCCGAACACATTGGGAAGGAATGTCACATTTGGTGAAGTTGTA 3261  
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006  
Db 3262 TTCACTATCCACCGACCGGACCTCCAGTGTCTTCCAGGATGAGCCTGGAGGTGAAG 3321  
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026  
Db 3322 AAGGCCAGAGCTGCTGTGGTGGCAGCAGTGGCTGTGGGAAGAGCAGAGTGGTCCAG 3381  
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046  
Db 3382 CTCCTGAGCGGTTCACGACCCCTTGGCAGGGAAGTGTCTTGTATGCAAGAAATA 3441  
Qy 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066  
Db 3442 AAGCGACTGAATGTCAGTGGCTCCGAGCACACCTGGGCATGCTGCCAGGAGCCCATC 3501  
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086  
Db 3502 CTGTTTGACTGACGATGCTGAGACATGCTATGGAGACACAGCCGGTGGTGTCA 3561  
Qy 1087 LeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106  
Db 3562 CAGGAAGAGATTGTGAGGCGCAGCAAGAGGCGCCAAACATACATGCTTCATCGAGTCACTG 3621  
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126  
Db 3622 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAAGTGTCTGCTGGTCCAGAA 3681  
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146  
Db 3682 CAACGATTTGCCATAGCTGCTGCTGCTGTGTAGACGCTCATATTTTGGTATGAA 3741  
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166  
Db 3742 GCCACGTCACTCTGGATACAGAAAGTGAAGAGTGTGTCCAAAGACCTTGGCAAGGCC 3801  
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186  
Db 3802 AGAAGAGCGCCACCTGCTGATTTGCTCACCCTGCTCCACCATCCAGATGCGAGC 3861  
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeu 1206  
Db 3862 TTATAGTGTGTTTCAGATGCGCAGAGTCAAGGACATGCGCAGCATCAGCAGTCTGCTG 3921  
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 3922 GCACAGAAAGGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 3963

## RESULT 6

5206352-3  
; Patent No. 5206352  
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,  
; Michael M.  
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA  
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
; NUMBER OF SEQUENCES: 4  
; CURRENT APPLICATION NUMBER: US/07/622,836  
; FILING DATE: 24-SEP-1990  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER: 892,575  
; FILING DATE: 01-AUG-1986  
; APPLICATION NUMBER: 845,610  
; FILING DATE: 28-MAR-1986  
; SEQ ID NO: 3  
; LENGTH: 4669  
5206352-3  
Alignment Scores:  
Pred. No.: 0 Length: 4669  
Score: 3519.50 Matches: 681

Percent Similarity: 73.84% Conservative: 245  
Best Local Similarity: 54.31% Mismatches: 265  
Query Match: 57.14% Indels: 63  
DB: 6 Gaps: 10  
US-09-873-409-5 (1-1222) x 5206352-3 (1-4669)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 575 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 634  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 635 GTGTTTGGAGAAATGACATATCTTTGCAATGCGAGGAATTTAGAGATCTGATGTCA 694  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 695 AACATCACTAATAAGAGATATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 754  
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 755 ATGACCAAGTATGCCATTATTTACAGTGAATTTGGTGGGGTGTGGTGTCTTAC 814  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 815 ATTCAAGTTTCAATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 875 TTTTTCATGCTATATATGCGCAGGAGATAGGCTGGTTTGTATGTGCACCATGTTGGGAG 934  
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 935 CTTAACACCCGACTTACAGATGATGCTCTAGATTAATCAAGTATTGGTGACAAAATT 994  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTCACTGGGTTTATAGTAGGATTACACGT 1054  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 1055 GGTGGAACTAACCTTGTGATTTTGGCCACGATGCTGTTCTGGACTGTGACGTGCT 1114  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 1115 GCTGGGCAAGACTATCTTCACTTATGATTAAGAACTCTTAGCGTATGCAAAAGCT 1174  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 1175 GGAGCAGTAGCTGAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTGGAGGACAA 1234  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 1235 AAGAAAGAACTTGAA----- 1249  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 1250 -----AGTACACAAAAATTTAGAGAA 1273  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTCTGCTTC 1333  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 1334 CTGCTGATCTATGATCTTATGCTGGGCTTCTGTGATGGGACCACTTGGTGGCTCTCA 1393  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1394 GGGGAA-----TATCTATTTGGAAGAAGTACTCACTGATTTCTTTCTGTTAATGGG 1447  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320



```
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAla 1046
DB 3668 CTCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTTGTATGCGAAGAATA 3727
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
DB 3728 AAGCGACTGAATGTTCACTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCATC 3787
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValPro 1086
DB 3788 CTCTTTGACTGCGACGATGCTGAGACATTCCTATGGAGACCAACAGCCGGGTGTCA 3847
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
DB 3848 CAGAGAGAGATCTGAGGCGAGCAAGAGGCGCCAAATACATACATCCCTTCATCGAGTCACTG 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
DB 3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACCTCAGCTCTCTGGTGCCAGAAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
DB 3968 CAACGCAATCCCATAGCTGCTGTCCTTTGTAGACAGCTCATATTTGCTTTTGGATGAA 4027
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
DB 4028 GCCAGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGCCCTGGACAAAGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
DB 4088 AGAAGAGCGCGCACCTGTCATTTGATTTGCTCAGCGCTGTCCACCATCCAGAGATGCAGAC 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnLysThrHisGlnGluLeu 1206
DB 4148 TTAATAGTGTGTTTTCAGATGGCAGATCAAGAGATGCGCAGCATCAGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
DB 4208 GCACAGAAAGGCATCTATTTTTCATGGTTCAGTGTCCAGGCT 4249

RESULT 7
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5984088
; GENERAL INFORMATION:
; APPLICANT: Mechnetr, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehrnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,447
; FILING DATE: 15-Nov-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3503.50 Matches: 677
Percent Similarity: 73.52% Conservative: 245
Best Local Similarity: 53.99% Mismatches: 269
Query Match: 56.88% Indels: 63
DB: 2 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-752-447-1 (1-4669)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACATCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTGGGAATGACATATCTTTGCAATGACGAAATTTAGACATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 695 AACATCACTAATAAGAGTGCATATCATGATGATACAGGCTTCTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuLeuPheGlyTyr 61
DB 755 ATGACCAGGTATGCTTATTATACAGTGGAAATGGTGGGGTGTCTGTTGCTGCTTAC 814
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTTTGGCTGGCAGCTGGAAGCAATATACAAAATTAGAAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGTTTGTATGTGCACGATGTTGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTTGGTCAAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACAGCT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTCTGCACTGTGAGTGT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1115 GTCTGGGCAAGATACATCTTCTTCAATTAAGATAAGAACTCTTAGCGTATGCAAAAGCT 1174
QY 181 GlyAlaValAlaGluGlnValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAATCTGTGATTCATTTGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
```







```
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 3368 GTCTTTGGTCCATGCGCGGGGCAAGTCAGTTCTTCTGACATGCAAGGCC 3427
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3428 AAAATATCAGCAGCCACATCATCATTCATTTGAAAAAACCCTTTGATTGACAGCTAC 3487
QY 967 SerGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3488 AGACGGAAGGCTTATGCGCAACACATTCGGAAGAAATGTCATTTGTTGAGTTGTA 3547
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3548 TTCAACTATCCACCCGACCGGACATCCAGTGTCTTCAGGACTGAGCTCGGAGGTGAAG 3607
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3608 AAGGCCAGACGCTGCTCTGTTGGGAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3667
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db 3668 CTCTGAGGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGTTGATGCCAAGAAATA 3727
QY 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3728 AAGCGACTGATGTTCTAGTGGCTCCGAGCACACTGGGCATCGTGTCCAGAGGCCATC 3787
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3788 CTGTTTCACTGCAGCATTTCTGAGACATTTGCTATGGACAAACAGCCGGTGGTGTC 3847
QY 1087 LeuAspGluLeuLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3848 CAGGAAGAGATCGTAGGCGGACGAAAGGCGCCAAACATACATGCTTCATCGAGTCACTG 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3908 CCTAATATATATAGCATTAAGTAGGAGACAAAGGAATCAGCTCTCTGTGGCCAGAAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3968 CAACGCTATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4027
QY 1147 AlaThrSerAlaLeuAspAsnSerGluLysValGlnHisAlaLeuAspLysAla 1166
Db 4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGGCCCTGGCAAGGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 4088 AGAGAAGCGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1206
Db 4148 TTAATAGTGTGTTTCAAGATGGCAGATCAAGAGCATGCGCAGCATCAGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCAAGTGTCCAGGCT 4249

RESULT 8
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechetrner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; USING IMMUNOLOGICAL AGENTS SPECIFIC FOR P-GLYCOPROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehrnen Hulbert & Berghoff Ltd.
```

```
STREET: 300 South Wacker Drive, Seventh Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/316,167
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/752,447
FILING DATE: 15-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: No. 6365357nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..424
FEATURE:
NAME/KEY: CDS
LOCATION: 425..4264
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 4265..4669
US-09-316-167-1
Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3503.50 Matches: 677
Percent Similarity: 73.52% Conservative: 245
Best Local Similarity: 53.98% Mismatches: 269
Query Match: 56.88% Indels: 63
DB: 4 Gaps: 10
US-09-873-409-5 (1-1222) x US-09-316-167-1 (1-4669)
QY 1 MetIleGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGTGGTGGGAACCTTGTGCTGCCATCATCCATGGGGTGGAGCTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAATGACATATCTTTGCAATGACAGAAATTTAGAGATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAAGTATATCAATGATACAGGCTTCTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAAGTATGCTTATTATTACAGTGAATTTGGTGGGTGCTGCTGCTGCTTAC 814
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGTTTCATTTTGGTGGCTGGCAGCTGGAGACAAATACACAAATAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
```





QY 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30  
Db 230 GTGTTCCGATACATGACAGATAGTTTACCAAGCAGAGACCGGCATTTCTGCGAGCGTT 289  
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45  
Db 290 ACTAATCAAGTGAATCAACAGTACACAGACCGTCCAGCAGCAGTCTGGAGGAGGAC 349  
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuLeuPheGlyTyr 61  
Db 350 ATGGCCATGPAACCCCTACTATTACACGGGCATTGGTCCGGTGTCTCATCGTTCCTAC 409  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 410 ATCCAGGTTTCACTTTGGTGTCTGGCAGCTGGGAGCAAAATACACAGATTAGCCCAAG 469  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 470 TTTTTCATGCCATCATGAATCAGGAGATAGGCTGGTTCACGTGAATGACGCTGGGGAG 529  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 530 CTCACACCCGGCTCACAGATGACGCTCCAAAATTAATCAGCGAATTGGTGACAACCT 589  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGCTTATTTGCGGTGTTTATAAGT 649  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 650 GGTGGAAGCTAACCTTGTAATTTGGCCGTCAGCCCTCTTATTTGGTGTGTCATCTGCC 709  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSerLysAla 180  
Db 710 ATGTGGCAAGGTACTGACTTCATTACTAATAAGAACTCCAGGCTTATGCGAAGCT 769  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATGGTGGTGGAGCAA 829  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 830 AAGAAGGAACCTGAA----- 844  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 845 -----AGGTACAATAAATAATTTAGAGAA 868  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCCAACATTTCCATAGGTATTGCCATC 928  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
Db 929 CTGTGTGCTATGCGTCTTATGCACTGGCAATCTGTGGTATGGGACCTCTTGGTCTCTCA 988  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 989 AATGAA-----TATTCTATGGACAAGTCTTACCGTCTCTCTCTATTATTATGGGG 1042  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1043 ACTTTCAGTATGGACATTTAGCCCCCAACATAGAACCTTTGCAATGCAAGAGGGGCA 1102  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1103 GCCTATGAAATCTTCAAGATAATTTGATAATGAGCCCAAGCATTTGACAGCTTCTCAACCAAG 1162  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysValSerPheAsnTyr 360  
Db 1163 GGACAAACACACAGATATAATGGGAAATTTGAAATTTAAATGTTTACTTCAACTAC 1222  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380

Db 1223 CCATCAGAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGTGAAGCGGCGAG 1282  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1283 ACGGTAGCCTGGTTGGCAACAGTGGCTGGGAAAGCACAACCTGTCAGCTGCTGCGAG 1342  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1343 AGGCTCTACGACCCCATAGAGGCGAGGTCAATATCGAGCAGCAGGACATCAGGACCATC 1402  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1403 AATGTGAGGTATCTCGGGAATCATTTGGGGTGGTCACTCAGGAAACCGTGTGTTGCC 1462  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
Db 1463 ACCACATGTCGCCAAACATTCGCTATGCGCGAGAAACGTCACCATGGATGAGATAGAG 1522  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProLeuLysPheAsn 480  
Db 1523 AAGCTGTCAAGGAAGCCAAATGCTATGACTTCATCATGAACCTGCCCAACAAATTAAC 1582  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1583 ACCCTGGTGGTAGAGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCCTTTGTTGATGAGGCCACGTCAGCTTG 1702  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1703 GACACAGAAAGCGAAGCCGCTGGTTTCAAGCGCTCTCGATAAGGCTAGAGAAGCCGACC 1762  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1763 ACCATTGTATGACTCACCGCTTGTACAGTGCGAATGCTGCACCTCATTTGCTGTTTT 1822  
QY 561 LysAspGlyMetLeuAlaGluLysGlyValHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1823 GATGGTGGTGCATTGTGGAGCAAGGAATCATCAAGAGCTCATGAAGAGAGAGGCAAT 1882  
QY 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600  
Db 1883 TACTTCAAACTTGTCTATGACACAG---ACTAGAGGAAATGAAATTTGAACCCAGAAATAAT 1939  
QY 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620  
Db 1940 GCTTATGAATCCCAAGTGACACTGGTGGCTCTGTAGTTGACTTTCAGAAAAATCA---AAA 1996  
QY 621 SerAspPheIleAspLysAla-----GluGlu 629  
Db 1997 TCTCCTTTAATAGGAGATCAATTCGCAAGATATCCACAGAGACAGACAGGAGAGA 2056  
QY 630 SerThrGlnSerLysGlu-----IleSerLeuProGluValSerLeuLeuLysIle 646  
Db 2057 AGACTTAGTTCGAAAGAGGATGTGGATGAAGATGTGCCTATGCTTTCCTTTTGGCAGATC 2116  
QY 647 LeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeu 666  
Db 2117 CTAAGCTAAATATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTATA 2176  
QY 667 AsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly 686  
Db 2177 AATGGGTGATCAACACAGGTTTGGCATAGTGTTCCTTCAAGGATTTAGGGGTTTTTCA 2236  
QY 687 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705  
Db 2237 AGAGACGACGACCATGAACCAACCAACGAATTTGTAATTTTCCCTTCTCTTCTG 2296  
QY 706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725

```
Db 2297 GTCATGGGAATGATTTCTTTTGTGTACGTACTCTTTCAAGGCTTCACATTTGGCAAGCT 2356
Qy 726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
Db 2357 GGAGAGATCTCTACCAAGCAGCCGATACATGTCTTCAATCCATGCTGCCAGAGAT 2416
Qy 746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
Db 2417 ATAAGCTGGTTTATGATACCAATAAAACACCACTGGCTCGCTGACTACCGCTCGCTAGT 2476
Qy 766 AspilealadlnleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
Db 2477 GAGCGCTTCTAATGTTAAAGGGGTATCGGCTCCAGGCTTCTGTAGTATCCCAAGATGTA 2536
Qy 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 805
Db 2537 GCAACCTTGGCAGCAAGAAATATCTTATCTTGTAGTCTATGCTGGCAGCTTACACTTTTA 2596
Qy 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaIleMetThr 825
Db 2597 CTGTGTAGTAATATATACCACTCATTTCTTCTGGTGGAAATATTGAAATGAAACTGTTGTCT 2656
Qy 826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
Db 2657 GGTCAAGCCTTGAAGGACAAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCA 2716
Qy 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
Db 2717 ATTGAAACTTCGCACTGTGTCTCTTGTAGTAGGAGCAGAGTTTGAAACTATGTAT 2776
Qy 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
Db 2777 GCCCAGAGCTTCGAGATACCATACAGAAATGCTTTGAAGAAAGACACAGCTTTGGATC 2836
Qy 886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
Db 2837 ACCTTCGCTTCAACCGGCGCATGATTTATTTTCTATGCTGCTGTTTCCGCTCGGT 2896
Qy 906 AlaTyrIleGlnAlaGlyArgMetThrProGluClyMetPheIleValPheThrAla 925
Db 2897 GCTTACTTGTGGCAGAGAACTCATGACGTTTGAATAATGTTATGTTGGTATTTCTGCT 2956
Qy 926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
Db 2957 GTTGTCTTGTGTGTCATGGCAGCAGGAAATACCATGTTTCATTCGCTCTGCTAGCGAAG 3016
Qy 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsnIleAspSer 965
Db 3017 GCCAAAGCTTCAGCATCCCATCATCAGGATCATTCAGAAAAATCCCGAGATTGACAGC 3076
Qy 966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
Db 3077 TACAGCAGGAGGCTTGAAGCCTTATGTTAGTGAAGAAATGTAATTTAATGGATC 3136
Qy 986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
Db 3137 ATGTTCAACTATCCACCGACCAACATCCAGTGTCTTCAGGAGCTGAGCTCGAGGTG 3196
Qy 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
Db 3197 AAGAAGGGGCAAGCGTTCGCTGGTGGCAGCAGTGGCTGGCGGAAGAGTACAGTGGTC 3256
Qy 1026 GlnLeuLeuGlnArgLeuTyrAspProValClnGlyGlnValLeuPheAspGlyValAsp 1045
Db 3257 CAGTGTCTGAGCGCTTCTACACCCCATCGCTGGTGAACAGATGTTCTAGATGGCAAGAA 3316
Qy 1046 AlaLysLeuLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
Db 3317 ATAAACAACCTCAACGTCAGTGGCTCGCGGCCACCTGGGCATTTGTCTCCAGGAGCCC 3376
Qy 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyLysAspAsnSerArgValVal 1085
Db 3377 ATCCTGTTTGACTGCAGCATCACCGAGAACATCGCCTACGAGACAAACACCGCTGTGCTG 3436
```

```
Qy 1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
Db 3437 TCTCATGAGGAGATCGTAGGCGCGCAGGAGGCCAACATCCACCATGTTTCAGCTCA 3496
Qy 1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGln 1125
Db 3497 CTGCTCTGAGAAATACAAACACCAAGTGGGAGCAAAAGGAGACTCAGCTGTGGGCGGCGAG 3556
Qy 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAsp 1145
Db 3557 AAGCAGCGATCGCCATCGCGCGCCCTCGTCAGACAGCTCACCATCTTACTTCTGGAT 3616
Qy 1146 GluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAlaLeuAspLys 1165
Db 3617 GAAGCCACATCAGCTCTGGATCAGGAGAGTGAAGAGTGTCCAGGAGCGCTGGACAAA 3676
Qy 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
Db 3677 GCCAGGAAAGCGCACCTGCTGATCGCGCACCGCTGTCCACCATCCAGAACGCA 3736
Qy 1186 AspLeuIleValValLeuHisArgGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
Db 3737 GACTTGTGTGTGATTCAGAACGCGCCAGGTCAAGAGACGCGCTGTCCACCATCCAGACGCTG 3796
Qy 1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db 3797 CTGCCCCAGAAAGCATCTATTCTCGATGGTTCAGGCT 3835
```

## RESULT 10

```
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; EARLIER FILING DATE: 1999-11-29
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1
```

## Alignment Scores:

Pred. No.:	0	Length:	4233
Score:	3494.00	Matches:	684
Percent Similarity:	72.87%	Conservative:	229
Best Local Similarity:	54.59%	Mismatches:	274
Query Match:	56.73%	Indels:	66
DB:	4	Gaps:	10

US-09-873-409-5 (1-1222) x US-09-450-105-1 (1-4233)

```
Qy 1 MetIleGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 170 ATGGCTCTGGAACTCTCGCTGCTATCATCCAGAACCTGCTTCCCTCTCGTATGCTG 229
Qy 21 ValLeuGlyLeuMetSerAspAsnLeuIle----- 30
Db 230 GTGTTCCGATACATGACAGATAGTTTTACCAAGCAGACCCGCAATTCGCGAGGTT 289
Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPheArgLeu----- 45
```

Db 290 ACTAATCAAGTGAATCAACAGTACACAGCCGTGAGCGACAGCAGTCTGGAGGAGGAC 349  
Qy 46 -----ThrLeuTyrTyrValGlyValGlyValAlaAlaLeuLeuPheGlyTyr 61  
Db 350 ATGGCCATGTACCCCTACTATTACACGGGCAATGGTGGCGGTGCTCATCTGTCCTAC 409  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgIleArgIleGln 81  
Db 410 ATCCAGGTTTCACTTTGGTGGCGGAGTGGGAGACAAATACACAAAGATTAGGCAGAG 469  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 470 TTTTTCATGCCATCATCAATCAGGAGATAGGCTGTTTGCATGAATGACGCTGGGAG 529  
Qy 102 LeuAsnThrArgMetThr--AspIleAspIleSerAspGlyIleGlyAspIleIle 120  
Db 530 CTCACACCCGGCTCACAGATGACGCTCCAAATTAATGACGGAATTTGGTGACAACCT 589  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys 140  
Db 590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTTATATAGGATTTATAAGT 649  
Qy 141 GlyTrpIleLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 650 GGTGGAAGCTAACCTTGTAATTTGGCGCTCAGCCCTCTTATTGGGTGTCATCTGCC 709  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysLeuLeuSerAlaTyrSerLysAla 180  
Db 710 ATGTGGCAAGGTACTGACTCTTATTACTTAATAAGAACTCCAGGCTTATGCGAAGCT 769  
Qy 181 GlyValAlaAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 770 GGAGCAGTTGCCAAGAGCTTTAGCAGCCATCAGAACTGTGATTGGTTGGAGGACAA 829  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 830 AAGAAGGAACCTTGAA----- 844  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 845 -----AGGTACAATAAATAATTTAGAGAA 868  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 869 GCTAAGAAGTTGGCATAAAGAAAGCCATCACCGCAACATTTCCATAGTATTGCCCTAC 928  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
Db 929 CTGTTGCTATCGCTCTTATGCACCTGGCATCTGGTATGGGACCTCTTGGTCTCTCA 988  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIleHisSer 300  
Db 989 AATGAA-----TATTCTATGGCAAGTGCTTACCGCTCTCTCTATTTATTGGGG 1042  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1043 ACTTTCAATTTGGACATTTAGCCCCCAACATAGAGCCTTTGCAATGCAAGAGGGCA 1102  
Qy 321 AlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThrAla 340  
Db 1103 GCCTATGMAATCTTCAAGATAATTAATGATGAGCAAGCATTCACAGCTTCTCAACCAAG 1162  
Qy 341 GlyTyrIleProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1163 GGACACAAACACAGACAGTATAATGGGAAATTTGGAAATTTAANAATTTTACTTCAACTAC 1222  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1223 CCATCAGAAAGTGAAGTTAAGATCTTGAAGGCTCAACTGAAGGTGAAGAGCGGGCAG 1282  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400

Db 1283 ACGGTAGCCCTGGTGGCAACAGTGGCTGTGGAAAGACAACTGTCTCAGCTGCTGCAG 1342  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1343 AGGCTCTACACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlnProValLeuPheGly 440  
Db 1403 AATGTGAGGTATCTCGGGAAATCATTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1462  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1463 ACCACGATTGGCGAAACATTCGTATGCCGCGGAAACGTCACCATGGATGATGATAGAG 1522  
Qy 461 ArgAlaAaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1523 AAAGCTGTCAAGGAAGCCAATGCTATGATCTTCATCATGAACCTGCCCCACAAATTAAC 1582  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1583 ACCCTGGTTGGTGAGAGAGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1643 GCCCGGCCCTGCTCGCAACCCCAAGATCTTTGTTGGATGAGGCCACGTCAGCCTTG 1702  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1703 GACACAGAAGCGGCGGTGTTGAGGCGCTCTGGATAAGGCTAGAGAGGCGCGACC 1762  
Qy 541 ThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1763 ACCATTGTATAGTCCCGCTTGTCTACAGTCGCAATGCTGACGCTCATTTGCTGTTT 1822  
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1823 GATGGTGGTGTCTGTGGAGCAAGGAATCATGAAGCTCATGAAGAGGAGGCAAT 1882  
Qy 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600  
Db 1883 TACTTCAAACTTGTCTATGACACAG---ACTAGAGGAATGAAATTCACACAGGAATAT 1939  
Qy 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620  
Db 1940 GCTTATGAATCCCAAGTCACACTGGTGGCTCTGACTTGACTTCAGAAAAATCA---AAA 1996  
Qy 621 SerAspPheIleAspLysAla-----GluGlu 629  
Db 1997 TCTCTTTAATAGGAGATCAATTCGCAAGATATCCAGAGAGATCCACAGAGACAGACAGGAGAGA 2056  
Qy 630 SerThrGlnSerLysGlu-----IleSerLeuProGluValSerLeuLeuLysIle 646  
Db 2057 AGACTTAGTTGCAAGAGGATGTGGATGAAGATGTCCTATGCTTTCCTTTTGGCAGATC 2116  
Qy 647 LeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeu 666  
Db 2117 CTAAGCTAAATATTAGTGAATGGCCCTATTAGTTGTGGGTGCTACTTTGTGCTGTATA 2176  
Qy 667 AsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly 686  
Db 2177 AATGGTGTCAATCAACCCAGTGTTCATAGTGTTCATGAAGATTGTAGGGGTTTTTCA 2236  
Qy 687 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705  
Db 2237 AGAGCAGACCATGAACCAACCAACGAATGTAACTGTTTTCCTTCTCTTCTCTG 2296  
Qy 706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725  
Db 2297 GTCATGGGAATGATTTCTTTTGTAGTACTTCTTTCAAGGCTTCACATTTGCAAGACT 2356  
Qy 726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745  
Db 2357 GGAGAGATCCTCACCAAGCGACTCCGATATCGTCTCTCAAAATCCATGCTGTCGACAGGAT 2416



```
QY 746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
D 2417 ATAAGCTGGTTTGTATGACCAATAAACCACACTGGCTCGCTGACTACCGAGCTCGTAGT 2476
QY 766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
D 2477 GAGCGCTTAAATGTTAAAGGGGCTATGGGCTCCAGGCTGCTGTAGTTATCCCAAGATGA 2536
QY 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleIleGlyTrpGluMetThrPheLeu 805
D 2537 GCAAACTTGGCACAGAAATATCTATCTTAGTCTATGGCTGGGAGCTTACACTTTTA 2596
QY 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
D 2597 CTGTAGTATATACCACTCATCTGTCTGGGTGGAATATTGAAATGAACCTGTTGCT 2656
QY 826 GlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAla 845
D 2657 GGTCAAGCCTTGAAGGACAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCA 2716
QY 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
D 2717 ATTGAATACTTCGCACTGTGTCTCTCTACCTAGGAGCAGAAAGTTTGAACATATGTAT 2776
QY 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
D 2777 GCCAGAGCTTGACATACCATACAGAAATGCTTGAAGAAAGCACACGCTTTGGGATC 2836
QY 886 CysTyAlaPheSerHisAlaPheIleTyPheAlaTyAlaAlaGlyPheArgPheGly 905
D 2837 ACCTTCGCTTACCAGGCCATGATTAATTTTCTCTATGCTGCTGTTTCCGGTTCGGT 2896
QY 906 AlaTyIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
D 2897 GCCTACTGTGTGCACGAGAACTCATGACCTTTGAAATGTTATGTTGTTATTTCTGCT 2956
QY 926 IleAlaTyGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTySerLys 945
D 2957 GTTGTCTTGTGTCATGGCAGCAGGAAATACCATTCCTCTCTGACTAGCGGAAG 3016
QY 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 965
D 3017 GCCAAAGCTCAGCATCCCATCATCAGGATCATTCAGAAATATCCCGAGATTGACAGC 3076
QY 966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
D 3077 TACAGCACGGAGGGCTTGAAGCCCTTAATTTGTTAAGGAAATGTGAATTTAATGAGATC 3136
QY 986 SerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSerIle 1005
D 3137 ATGTTCAACTATCCACCCAGCCCAACATCCAGTGTCTCAGGAGCTGAGCTTCGAGGTG 3196
QY 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
D 3197 AAGAAGGGCAAAAGCTTCGCTGTGGCAGCAGTGGCTCGCGGAAGAGTACAGTGGTC 3256
QY 1026 GlnLeuGlnArgLeuTyArgProValGlnGlnValLeuPheAspGlyValAsp 1045
D 3257 CAGCTGCTCAGCGCTTCTCAACCCCATCGCTGGAACAGTGTTCAGATGCGCAAGAA 3316
QY 1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
D 3317 ATAACAACCTCAACGTCAGTGGCTCGGCCCACTCGGCGCATTTGTCCAGGAGCCC 3376
QY 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyArgGlyAspAsnSerArgVal 1085
D 3377 ATCCTGTTTGCATCGACATCACCGCAACATCGCCTACCGAGACAAACAGCCGTGCTGTG 3436
QY 1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGly 1105
D 3437 TCTCATGAGGAGATCGTGAGGGCCGCCAGGAGGCCAACATCCACCAAGTTTCATCGACTCA 3496
```

```
QY 1106 LeuProGluLysTyAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
D 3497 CTGCTGTAGAAATACAAACACAGAGTGGGAGACAAAGGACTCAGCTGTGGGGCGGCGAG 3556
QY 1126 LysGlnArgLeuAlaIleAlaAargAlaLeuGlnLysProLysIleLeuLeuAsp 1145
D 3557 AAGCAGCGCATCGCATCGCGCGCCCTCGTCAGACAGCCCTACATCTTACTTCTGGAT 3616
QY 1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
D 3617 GAAGCGACATCAGCTCTGATACGGAGAGTGAAGAGTCTCCAGGAAGCGCTGGACAAA 3676
QY 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
D 3677 GCCAGGGAAGCGGACCTGTCATTTGTATCGCGCACCGGCTGTCCACCATCCAGAACGCA 3736
QY 1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
D 3737 GACTTGTATCGTGTGATTTCAGAACCGGCCAGGTCAAGGACGCGGCCCCACCCAGCAGCTG 3796
QY 1206 LeuArgAsnArgAspIleTyPheLysLeuValAsnAla 1218
D 3797 CTGCCCCAGAAAGGCACTATTCTCGATGTTTCAGGCT 3835

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DW4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 bases
; TYPE: nucleic acid
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.: 0 Length: 4669
Score: 3473.50 Matches: 673
Percent Similarity: 73.34% Conservative: 246
Best Local Similarity: 53.71% Mismatches: 271
```



Query Match: 56.40% Indels: 63  
DB: 2 Gaps: 10

US-09-873-409-5 (1-1222) x US-08-583-276-18 (1-4669)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20  
Db 575 ATGGTGGTGGAACTTGGCTGCATCATCCATCGGGCTGGACTTCTCTCATGATGCTG 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAATTTAGAAAGATCTGATGCTCA 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTy:Ser:PhePhe----- 43  
Db 695 AACATCACTAATAAGAAATGATATCAATGATACAGGGTCTTCTCAATGAACTTGGAGGAAGAC 754

Qy 44 -----ArgLeuThrLeuTyTyValGlyIleGlyValaAlaLeuIlePheGlyTy 61  
Db 755 ATGACAGGATATGCCATTATTTACAGTGAATTTGGTCTGGGTGCTGGTCTGCTTAC 814

Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleA:GlyGln 81  
Db 815 ATTCAAGTTTCATTGGTCCCTGGCAGCTGGAAGCAAAATACACAAATTAGAAAACAG 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 875 TTTTTCATGCTATAATCGCACAGGAGATAGGCTGTTGATGTCACGATGTTGGGCAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 935 CTTAAACCCCGCTTACAGATGATGCTCTAAGATTAAATGAATTTGTTGTGACAAAATT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 995 GGAATGTTCTTCAGTCATGGAACAACTTTTCTACCTGGGTTTATAGTAACTTACAGCT 1054

Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThr:SerProLeuIleMetAlaSerAlaAla 160  
Db 1055 GGTGGAGCTAACCTTGTGATTTGGCCATCAGTCTCTTCTGGCAGTCTCAGCTGCT 1114

Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180  
Db 1115 GTCTGGGCAAGATATCTATCTTCTTACTATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 1175 GGAAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 1234

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTy:Phe 220  
Db 1235 AAGAAAGAACTTGAA----- 1249

Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyThrGlnAsnLeuLysAsp 240  
Db 1250 -----AGGTACACACAAAATTTAGAGAA 1273

Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260  
Db 1274 GCTAAAGAAATTTGGGATAAAGAGCTTATTACAGCCAAATTTCTATAGGTGCTGCTTTC 1333

Qy 261 PhePheMetAsnGlyTyThrTyGlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsn 280  
Db 1334 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGGTATGGGACCACCTTGGTCTCTCA 1393

Qy 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1394 GGGGAA-----TATTCTATTGGACAAAGTACTCACTGATTTCTTTCTGTTAATAATTGGG 1447

Qy 301 SerTy:CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1448 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGAGCA 1507

Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1508 GCTTATGAAATCTTCAAGATAATTAAGTAAGCCCAAGTATTGCACAGCTATTGCAAGAGT 1567

Qy 341 GlyTy:LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTy 360  
Db 1568 GGGCACAACACAGATAATTAATTAAGGGAATTTGGAATTCAGAATGTTCACTTCAGTTAC 1627

Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1628 CCATCTCGAAAAAGATTAAAGATCTTGAAGGGCTCAACTGAAGCTGCAGAGTGGGCAG 1687

Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGln 400  
Db 1688 ACGTCCCCCTGGTTGGAAACAGTGGCTGTGGAAAGACCAACACAGTCCAGCTGATGTCAG 1747

Qy 401 ArgLeuTy:AspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGTCAGCAGGATATTAGACCATA 1807

Qy 421 AsnValArgHisTy:ArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1808 AATGTAAGGTTTCTACGGGAATCAATGGTGTGGTCAGTCAGGAACCTGATTTGTTGCC 1867

Qy 441 ThrThrIleSerAsnAsnIleLysTy:ArgAspAspValThrAspGluGluMetGlu 460  
Db 1868 ACCACGATGCTGAAACAACTTCGCTATGGCCGTGAAATGTCAACATGATGATGATGAG 1927

Qy 461 ArgAlaAlaArgGluAlaAsnAlaTy:AspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1928 AAGCTGTCAAGAACCCAAATGCCTATGATCTTATCATGAACCTGCTCATATAATTGAC 1987

Qy 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1988 ACCCTGTTGGAGAGAGAGGGCCCGCTGAGTGTGGTGGCAGAGCAGAGGATCGCCATT 2047

Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 2048 GCACGTGCCCTGGTTGCAACCCCAAGATCTCTCTGCTGATGAGGCCACGTCAGCCTTG 2107

Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 2108 GACACAAAGCCGACAGCTGTTTCAAGTGGCTCTCGATAAGGCCAGAAAGTCGGACC 2167

Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 2168 ACCATTGTGATGCTCATCGTTGTCTACAGTTCGTAATGTCAGCTCGCTCGTGGTTTC 2227

Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 2228 GATGATGGAGTCACTTGGGAAAGGAATCATGATGAACCTCATGAAGAGAAAGGCATT 2287

Qy 581 TyTy:SerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
Db 2288 TACTTCAAACTTGTCAATGACAGACAGCAGGAAATGAAGTTGAATTTAGAAAATGACGCT 2347

Qy 595 GluGlnMetGlnSerMetThrTy:SerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 2348 GATGAATCCAAAGTGAATTTGATGCTTGGAAATGCTCTCAATGATTCAGATCACGCT 2407

Qy 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
Db 2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAACCCCAAGACAGAAAG 2467

Qy 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647  
Db 2468 CTTAGTACCAAGAGAGCTCTGGATGAAGATATACCTCCAGCTTTCTTTGGAGGATATG 2527

Qy 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
Db 2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2587

Qy 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687

```
Db 2588 GGAGGCGTGCACACGATTTGCAATAATATTTTCAAGAGATTAGGGTTTITACAAGA 2647
Qy 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db 2648 ATTGATGATCCTGAACAAACAGACAGAAATACATACTTTTTCATATTGTTTTCAGCC 2707
Qy 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db 2708 CTTGGAAATTAATTTTATATACATTTTCTTCAGGTTTCACATTTGGCAAGCTGGA 2767
Qy 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db 2768 GAGATCCTCACCAGGCGCTCCGATACATCGTTTTCGATCCATGCTCAGACAGGATGTG 2827
Qy 747 AlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAsp 766
Db 2828 AGTTGGTTTGTATGACCCCTAAACACACACTGGAGCATTTGACTACCGGCTCGGCAATGAT 2887
Qy 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db 2888 CTTGCTCAAGTTAAAGGGGTATAGTTCCAGGCTGCTGTATTAATCCCAAGATATAGCA 2947
Qy 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db 2948 ATTCTTGGGACAGGAATAATATATCTTCTCATCTGTTGGCAACTTAACACTGTTACTC 3007
Qy 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 3008 TTAGCAATTTGATCCCACTTCATTCGAATAGAGGAGTGTGTAATGAATAAATGTTGTCTGGA 3067
Qy 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 3068 CAAGCACTGAAGATAAGAAAGAACTAGAGGTGCTGGGAAGATCGCTACTGAGCAATA 3127
Qy 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 3128 GAAACACTTCGGAACCGTGTCTTCTGACTCAGGAGCAGAAGTTTGAACATATGATGCT 3187
Qy 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 3188 CAGAGTTTGCAGGTACCATCAGAAACTCTTTGAGGAAGCACACATCTTTTGAATPACA 3247
Qy 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 3248 TTTTCTTCCACCGCAAGATGATATTTTCTATGCTGGATGTTTCCGGTTTGAGCC 3307
Qy 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 3308 TACTTGTGGCACATAAACTAATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTT 3367
Qy 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 3368 GTCTTTGTGTCCTATGGCGGTGGGCAAGTTCATTTCTCTGACTATGCCAAGCC 3427
Qy 947 LysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArg 966
Db 3428 AAAATATCAGAGCCCAACATCATCATGATCATTTGAAAAAACCCTTTGATTTGACAGCTAC 3487
Qy 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3488 AGCAGGAGGCGCTAATGCCGAACACATTTGGAGGAAATGTCACTTTGGTGAATGTA 3547
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3548 TTAACATATATCCACCCGACCGGACATCCAGTCCCTTCAGGACTGAGCCCTGGAGGTGAAG 3607
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3608 AAGGCCAGACGCTGGCTCTGGTGGGACAGATGGCTGTGGGAAGACACAGTGGTCCAG 3667
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
```

```
Db 3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGTGATGGCAAGAAATA 3727
Qy 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3728 AAGCGACTGAATGTTTCAGTGGCTTCGAGCACACCTGGGATCGTGTCCAGGNGCCCATC 3787
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3788 CTGTTTGTACTGCAGCATTCCTGAGAACATTTGCTATGGACACACAGCCGGTGGTGTC 3847
Qy 1087 LeuAspGluIleLysGluAlaAlaAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3848 CAGGAAGAGATCGTGGGCGCAGCAAGAGGAGGCAACATACATGCTTCATCGAGTCACTG 3907
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3908 CTTAATAATATAGCTAATAGTAGGAGCAAGAGAACTAGCTCTCTGTTGGGCCAGAAA 3967
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146
Db 3968 CAACGCAATGGCATAGCTCGTGCCTTGTAGACAGCCCTCATATTTGCTTTTGGATGA 4027
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAAAGAACCCCTGGACAAACC 4087
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 4088 AGAAGAGCGCGACCTGTCATTTGATGCTCAGCGCTGTCCACCATCCAGATGGAGAC 4147
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1206
Db 4148 TTAATAGTGGTGTTCAGAAATGGCAGAGTCAAGCAGCATGGCAGCATCAGCAGCTGCTG 4207
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 1219
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTGCAGTGTCCAG 4246
```

## RESULT 12

```
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,823
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/168,621
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,433
; FILING DATE: 27-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Karta, Glenn E.
; REGISTRATION NUMBER: 30,649
```

REFERENCE/DOCKET NUMBER: PA-0012 CIP 2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301 527-2058  
 TELEFAX: 301 208-6997  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2726 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-461-823-1

## Alignment Scores:

Pred. No.: 1,31e-245 Length: 2726  
 Score: 2280.00 Matches: 435  
 Percent Similarity: 76.55% Conservative: 159  
 Best Local Similarity: 56.06% Mismatches: 168  
 Query Match: 37.02% Indels: 14  
 DB: 1 Gaps: 4

US-09-873-409-5 (1-1222) x US-08-461-823-1 (1-2726)

QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
 Db 2 ATTGAGAAAGCTGTCAAGGAGCCCAATGCTATGACTTTATCATGAACCTGCCTCATAAA 61  
 QY 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
 Db 62 TTTGACACCTGTTGGAGAGAGAGGGGCCAGTTCAGTGGTGGCAGAGCAGAGGATC 121  
 QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleIleLeuAspGluAlaThrSer 518  
 Db 122 GCATTTGCAGCTGCCCTGGTTCGCAACCCCAAGATCTCTGCTGGATGAGGCACGTC 181  
 QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538  
 Db 182 GCCTTGACACAGAAAGCGAAGCAGTGGTTCAGTGGCTCTGGATAAGCCAGAAAGGT 241  
 QY 539 ArgThrThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558  
 Db 242 CGGACCAACCATTTGTGATGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCG 301  
 QY 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578  
 Db 302 GGTTCGATGATGGAGTCAATGTGGAGAAAGGAAATCATGATCAATCATGAAAGAGAAA 361  
 QY 579 GlyLeuTyrTyrSerLeuValMetSerGln-----AspIleLysLys 592  
 Db 362 GGCATTCTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAT 421  
 QY 593 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 612  
 Db 422 GCAGCTGATGAATCCAAAGTGAATGATGCTTGGAAATGCTTCAAATGATTCAAGA 481  
 QY 613 LeuHisSerVal-----LysSerIleLys---SerAspPheIleAsp 625  
 Db 482 TCCAGTCTTAATAGAAAAAGATCAACTCGTAGGAGTGTCCGCGATCACAAAGCCCAAGAC 541  
 QY 626 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysLys 645  
 Db 542 AGAAACCTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCCTTTGGAGG 601  
 QY 646 IleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerVal 665  
 Db 602 ATTATCAAGCTAAATTTAACTGAATGGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 661  
 QY 666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 695  
 Db 662 ATAAATGGAGGCGCTGCAACACGACATTTGCAATTAATATTTTCAAAGATTTATAGGGTTT 721

QY 686 GlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704  
 Db 722 ACAAGAAATTTGATCCTGAAACAAAACGACAGAAATAGTAACATTTGTTTCACTATTGTTT 781  
 QY 705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724  
 Db 782 CTAGCCCTTGAATATTCTTTTATTATCATTTTCCITTCAGGGTTTCACATTTGGCAAA 841  
 QY 725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744  
 Db 842 GCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTCCGATCCATGCTCAGACAG 901  
 QY 745 AspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764  
 Db 902 GATGTGAGTGGTTTGTATGATGACCTTAAAAACACCACTGGAGCATTTGACTACCGGCTCGCC 961  
 QY 765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784  
 Db 962 AATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCAGGCTTCTGCTAATATCCCAAGAT 1021  
 QY 785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPhe 804  
 Db 1022 ATAGCAAAATCTTGGGACAGGAATATATATCTTCATCTATGTTGGCACTAACACTG 1081  
 QY 805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824  
 Db 1082 TTACTCTTAGCAATTTGATCCCATCTATGCAATAGCAGGAGTGTGTTGAAATGAAATGTTG 1141  
 QY 825 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844  
 Db 1142 TCTGGCAAGCACTGAAGATAAGAAAGAACTAGAAGTGTCTGGGAAGATCGCTACTGAA 1201  
 QY 845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864  
 Db 1202 GCATAGAAAATCTCGAACCGTGTGTTCTTTGACTCAGCAGCAAGTTGTAACATATG 1261  
 QY 865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleLeuGly 884  
 Db 1262 TATGCTCAGAGTTTGCAGGTACCATACAGAAATCTTTGAGGAAGACACATCTTTGGA 1321  
 QY 885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904  
 Db 1322 ATTACATTTTCTTCCACCGCAATGATGATATTTTCTATGCTGGATGTTTCCGGTTT 1381  
 QY 905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924  
 Db 1382 GAGGCTACTTGGTGGCACAATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTTCA 1441  
 QY 925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944  
 Db 1442 GCTGTTGCTTGTGGTGCATGCGCGGGGCAAGTCAGTTCTCTGCTGCTATGCTG 1501  
 QY 945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964  
 Db 1502 AAAGCCAAATATCAGCAGCCACATCATCATGATCATTTGAAATAAAACCCCTTTGATG 1561  
 QY 965 SerArgSerGlnGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984  
 Db 1562 AGTACAGCAGGAGGCTTAATGCCGAACACATTTGGAAGGAATGTCACTTTGTTGTA 1621  
 QY 985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004  
 Db 1622 GTTGTATTCAACTATCCCAACCGCAGCATCCAGTCTTTCAGGACTGAGCCTGGAG 1681  
 QY 1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024  
 Db 1682 GTGAAGAGGGCAGACGCTGCTGTGTGGCAGCAGTGGCTGTGGGAAGACACAGATG 1741  
 QY 1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyVal 1044  
 Db 1742 GTCCAGCTCTCGAGCGGTTCTACGACCCCTTGGCAGGGAAGTGTCTGTGTTGCAAA 1801  
 QY 1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1

Alignment Scores:
Pred. No.: 6,01e-243 Length: 4002
Score: 2259.00 Matches: 494
Percent Similarity: 58.41% Conservative: 232
Best Local Similarity: 39.74% Mismatches: 446
Query Match: 36.68% Indels: 72
DB: 2 Gaps: 12

US-09-873-409-5 (1-1222) x US-08-996-545-1 (1-4002)
QY 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db 368 GTGCCATTGCTGCGCGTGCAGCTTCCAGA-----GGATAATGTTAT 409
QY 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrVal 50
Db 410 ATCAATCTCGTAGAC-CAGTTCTATGATGAATTGACCAAGACGTACTGTACTCGTA 468
QY 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
Db 469 TACCTCGGTATCGCGGAGTTGTGCTACTGTCTATGTTAGTACTGTTGGCTTCATATACC 528
QY 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db 529 GGAGAACACGCCACGACGAAGATCCGCGAGTATCTCTGAGTCTATCTCCCGCAAC 588
QY 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db 589 ATTGGCTATTTTGATAAACTCGTGCCTGGGGAAGTGACCCCGTATAACAGCCGATACA 648
QY 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db 649 AACCTTATCCAGGATGCGCATTTCCGGAGAAGGTGGTCTCACTTTGACTGCGCTGGCGACA 708
QY 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db 709 TTCGTGACGACATTCATTATTCGCTACGTCAATATCTGGAAGTTGGCTCTAATTTGCAGC 768
QY 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
Db 769 TCAACAAATGTGCGCCCTCGTCTCACCATGGCGGGTGTCTCAGTTTATCATCAAGTAC 828
QY 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluValLeuSer 189
Db 829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGCGCGGCGACCTGTTGCGGAAGAGGTATCAGC 888
QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
Db 889 TCCATCAGAAATGCCACGCGTTTGGCACCAAGAGCTTGGCAAG-----936
QY 210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
Db 936 -----936
QY 230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
Db 937 -----CAGTATCAGGTGCCATTTAGACGAAGCTTAGAAATGGGAAACAAAGAACCCAG 987
QY 250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
Db 988 ATTGTCATGGGTTTCATGATTGGCGCCCTTATTTGGCCCTTATGTCGAACACTACGGTCTT 1047

RESULT 13
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1

Alignment Scores:
Pred. No.: 6,01e-243 Length: 4002
Score: 2259.00 Matches: 494
Percent Similarity: 58.41% Conservative: 232
Best Local Similarity: 39.74% Mismatches: 446
Query Match: 36.68% Indels: 72
DB: 2 Gaps: 12

US-09-873-409-5 (1-1222) x US-08-996-545-1 (1-4002)
QY 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db 368 GTGCCATTGCTGCGCGTGCAGCTTCCAGA-----GGATAATGTTAT 409
QY 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrVal 50
Db 410 ATCAATCTCGTAGAC-CAGTTCTATGATGAATTGACCAAGACGTACTGTACTCGTA 468
QY 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
Db 469 TACCTCGGTATCGCGGAGTTGTGCTACTGTCTATGTTAGTACTGTTGGCTTCATATACC 528
QY 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db 529 GGAGAACACGCCACGACGAAGATCCGCGAGTATCTCTGAGTCTATCTCCCGCAAC 588
QY 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db 589 ATTGGCTATTTTGATAAACTCGTGCCTGGGGAAGTGACCCCGTATAACAGCCGATACA 648
QY 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db 649 AACCTTATCCAGGATGCGCATTTCCGGAGAAGGTGGTCTCACTTTGACTGCGCTGGCGACA 708
QY 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db 709 TTCGTGACGACATTCATTATTCGCTACGTCAATATCTGGAAGTTGGCTCTAATTTGCAGC 768
QY 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
Db 769 TCAACAAATGTGCGCCCTCGTCTCACCATGGCGGGTGTCTCAGTTTATCATCAAGTAC 828
QY 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluValLeuSer 189
Db 829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGCGCGGCGACCTGTTGCGGAAGAGGTATCAGC 888
QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
Db 889 TCCATCAGAAATGCCACGCGTTTGGCACCAAGAGCTTGGCAAG-----936
QY 210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
Db 936 -----936
QY 230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
Db 937 -----CAGTATCAGGTGCCATTTAGACGAAGCTTAGAAATGGGAAACAAAGAACCCAG 987
QY 250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
Db 988 ATTGTCATGGGTTTCATGATTGGCGCCCTTATTTGGCCCTTATGTCGAACACTACGGTCTT 1047
```



```
Db 3202 CGAAGCCACAAATGTGATACTGGTCTGAAGAGGGCGAGAGGCTCGAAGCGTGGAGGTT 3261
Qy 979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
Db 3262 GAATCGAATTTAGAACGTGCATTCAGATACCGACCGCCCGCAGAACAGCTGTCTG 3321
Qy 999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
Db 3322 CGCGGCTTGACCTGACCGTGAAGCCTGGACAATATGTCGGCTTGTGCGACCCAGCGGT 3381
Qy 1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
Db 3382 TGTGGCAAGATACCACTGCTATTCCTTTGAGCGCTTTTACGATGCGATTCGCGGTCC 3441
Qy 1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
Db 3442 ATCTTCTTGTATGGAGGACATAGTAACTAAATATCACTCTACCGACGCTTCTG 3501
Qy 1059 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
Db 3502 TCATGTGTACGAGGAGCGGACACTGTACCGGCGCACCATCAAGGAAACATCTTACTT 3561
Qy 1079 GlyAspAsnSerArgValValProLeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsn 1098
Db 3562 GGTATTCTGAAGATGACGTACCGGAAGATTTCTGATTAGGCTTGCAAGGACGCTAAT 3621
Qy 1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
Db 3622 ATTCACACTTCATCATGTCTCCCGAGGGCTTTAATACAGTTGTGCGACGAGGGA 3681
Qy 1119 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
Db 3682 GGCATGTGTGTGCGCGCCAAAGACGCTGTGSCCATTTGCCGAGCCCTTCTTCGGGAT 3741
Qy 1139 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1158
Db 3742 CCCAAATTCCTTCTTCGTGATGAGCGACGCTCAGCCCTCGACTCCGAGTCAGAAAAGGTC 3801
Qy 1159 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1178
Db 3802 GTCCAGCGCGCTTGGATGCGCTGCGCGAGCGCCGAAACCAATCGCGCTGTGCACACCGA 3861
Qy 1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1198
Db 3862 CTCAGCACGATTCAAAGGCGGAGTTATCTATGTTTTCGACCAAGGCAAGATCGTCAA 3921
Qy 1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db 3922 AGCGGAACGACACGCAACTGGTCCAGAAAAGGCGCGGTACTACGAGCTGTCAACTTG 3981
Qy 1219 GlnSerVal 1221
Db 3982 CAGAGCTTG 3990
```

## RESULT 14

```
US-08-996-545-3
; Sequence 3: Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,545
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4002 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:
Pred. No.: 6,01e-243 Length: 4002
Score: 2259.00 Matches: 494
Percent Similarity: 58.41% Conservative: 232
Best Local Similarity: 39.74% Mismatches: 446
Query Match: 36.68% Indels: 72
DB: 2 Gaps: 12

US-09-873-409-5 (1-1222) x US-08-996-545-3 (1-4002)
Qy 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db 368 GUGCCAUUGCGCGGCGUCGCUUCCAGCA-----GGAAUUAUGUUAU 409
Qy 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrVal 50
Db 410 AUCAAAUUCUGUACGAC-GAGUUCUAUGAUAUACCAAGAACGACGUAUCUUAUCC 468
Qy 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleThr 70
Db 469 UACUCUGUAUUGCGGAGUUGUCACUGCUAUGUUAUAGUACUGUUGGCUUUAUUAUCC 528
Qy 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db 529 GGAAGACACGCCACGACGAGAACUCCGCGUAUUAUACUUGACUUAUCCGCGCCAGAAC 588
Qy 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db 589 AUUGGCUAUUUUGAUAACUCUGGUGCGGGAAGUACCAACCGUAUAACAGCGCAUACA 648
Qy 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db 649 AACCUUAUCCGAGGAGCAUUCGGAAGGUGGUGUCUUAUUAUUAUUAUUAUUAUUAU 708
Qy 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db 709 UUGUCACAGCAUUAUUAUCCGCUACGUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 768
Qy 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
Db 769 UCAACAUAUUGUGCCUUCUUCACCAUGGGGCGGUGUUCUCAGUUAUUAUUAUUAUUAU 828
Qy 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluValLeuSer 189
Db 829 AGCAAAAAGUUGCUUGACGACGUAUCCGUGAGCGGCGACUGUUGCGGAGGAGGUAUCC 888
```



Db	1882	ACUAUUGUGAUCGCUCAUCGCCUUCUCCAGCAUCAAAACGGCGCAACAACUUGUGUUCUC	1941
Qy	561	LYeAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu	580
Db	1942	GUCAUUGCGCAAAAUUCUGCAACAAGGACUCACACGAUAAUUGGUUAGCCGCGGAGCGCU	2001
Qy	581	TyTySerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet	600
Db	2002	UAUCGCAAAACUUGUGAGGCGUACAGCUAUAAGAACAGAAAGGAGUGAGCGCCUUGAG	2061
Qy	601	ThrTySerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValLysSer	618
Db	2062	GAGCCGACGCGAGGAUCUCACGAUAGCAGAUUUGCCAAAUCAAAACUCGCGUCAAGC	2121
Qy	619	IleLysSerAspPhe-----IleAspLys-----AlaGluGluSer	630
Db	2122	GCAUCAUCGACUUCGACGGAAAAACCCACCAUUGACCGCAGCGGCACCCACCAAGUCU	2181
Qy	631	ThrGlnSerLysGluLeuSer-----LeuProGluValSerLeuLeu	644
Db	2182	GUUUCAGCGCGAUUCUUUUAAGAGACCCCGGAAACAAACUCCGAAAUACUAUAUUG	2241
Qy	645	LysIleLeuLys-----LeuAsnLysProGluTyrProPheValValLeuGly	660
Db	2242	ACGCGUCUCAAAUUGUUGUCUUCUUAACCGCCCGGAAUCCCGACUAGCUCUACUGGU	2301
Qy	661	ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys	680
Db	2302	CUUGUCUUCACUGUUGUUGAGCUGUGUGGCGCCAAACCAACGACGACGUCUAUUGCUAA	2361
Qy	681	IleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGlu	698
Db	2362	GCCAUCGACACACUCUCGCGCCAGAAUACAUAUAGAACGACGUCGACUAGUCGGAU	2421
Qy	699	IleTySerMetIlePheValIleLeuGlyValIleCysPheValSerTyTyPheMetGln	718
Db	2422	UUCUGUCAUUGAUGUUCUUGUGUGUGUAUCUUCAGUUUAUACACGACGACUACAAU	2481
Qy	719	GlyLeuPheTyTyGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe	738
Db	2482	GGUCCGCAUUGCCGUAUGUCCGAGACAUUUAUUGCGCGGAGAGACGACUCCGUU	2541
Qy	739	LysAlaMetLeuTyTyGlnAspIleAlaTyPheAspGluLysGluAsnSerThrGlyGly	758
Db	2542	CGGACGAUACUCCGUCACAGACAUUGCUUUAUACAGGAAGAGAAUAGACCGCGCGCU	2601
Qy	759	LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle	778
Db	2602	CUGACCUUUCUGUCCACCGAGACGAAGCAUCUCUCGCGUUAUAGCGGUGAGACUCA	2661
Qy	779	GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyTy	798
Db	2662	GGCACGAUUGAUGACCUCCACGACCCUAGAGCGGCUUAUUAUUGCCCGCGCGAUU	2721
Qy	799	GlyTyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet	818
Db	2722	GGUGGAAAUUGCCUUAUGUUAUCUGUUGUGCGGUUCCUCCUGGCAUGCGGUUC	2781
Qy	819	IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla	838
Db	2782	UACCGAUUUAUUGUAGUAGCCCAAGUUCAUACACGCUCAAGCUUGCUUUAUAGAGGAU	2841
Qy	839	GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu	858
Db	2842	GCAAAACUUCUGCGGAGGCUACUGCUUAUCCGACAGUUGCGUUCUAUAACCCGGGAA	2901
Qy	859	LysAlaPheGluGlnMetTyTyGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys	878
Db	2902	AGGAGUCUGGAGAAUUAUACCAUGCCAGCCAGCUAGCGACGACGAGGACGACUUAUC	2961
Qy	879	LysAlaGlnIleIleGlySerCysTyTyAlaPheSerHisAlaPheIleTyTyPheAlaTy	898





Db 529 GGAAGACACGCCACGAGAGATCCGCGAGTATTACTTGAGTCTATCTCGCCGACGAGAC 588  
Qy 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109  
Db 589 ATTGGCTATTTTGATTAACCTCGGTGCGGGGAGGTGACCAACCGGTATACAGCCGATACA 648  
Qy 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129  
Db 649 AACCTTATCCAGGATGGCATTTTCGAGAGAGGTGCGTCTCACTTTGATCGCCCTGGCGACA 708  
Qy 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149  
Db 709 TTCGTACAGCATTCATTATTCGCTACGTACGTAATCTGGAAGTTGGCTCTAATTTCCAGC 768  
Qy 150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169  
Db 769 TCAACAATTGTGSCCCTCGTTCTCACCATGGGCGGTGTTCTCAGTTTATCATCAAGTAC 828  
Qy 170 ThrSerLysGluLeuSerLysAlaValAlaValAlaValAlaValGluValLeuSer 189  
Db 829 AGCAAAAGTTCGCTTGACAGCTACCGTGCAGCGCGGCACTGTTTCGGAAGAGGTTCATCAGC 888  
Qy 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209  
Db 889 TCCATCAGAAATGCCACAGGTTTGGCCCAAGACAAGCTTCGGAAG----- 936  
Qy 210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229  
Db 936 ----- 936  
Qy 230 \*\*\*PheValargTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249  
Db 937 -----CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCCAG 987  
Qy 250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269  
Db 988 ATGTGTCATGGTTTCATGATGGCGCATGTTTGGCCTTATGACTCGAACTACGGTCTT 1047  
Qy 270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289  
Db 1048 GGTCTTCGATGGGTCTCGTTTCTCGGTAGATGGT-----CGAGTCGATGGGTGAT 1101  
Qy 290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309  
Db 1102 ATTCTCACAGTCTCATGGCATCTTGATCGGATCGTTCTCTCGGGAACGTTAGTCCA 1161  
Qy 310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329  
Db 1162 AATGCTCAAGCATTTACAACGCTGTGGCGCGCGCCGCAAGATATTTGGAACGATCGAT 1221  
Qy 330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349  
Db 1222 CGCAGTCCCATAGATTCATATTCGACGAAGGGAAGAGCGCTCGACCATTTTGGGGC 1281  
Qy 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369  
Db 1282 CACATTGAGTTACGCAATGTCAAGCATATTTACCATCTAGACCCGAGGTACCGGTCA 1341  
Qy 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389  
Db 1342 GAGGATGTTTCTCTGTCATCGCCGCTGGAAAAACAACCGCTTTAGTCGCGCCCTCTGCG 1401  
Qy 390 SerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAspGlyPhe 409  
Db 1402 TCTGGAAGAGTACGGTGGTGGCTTGTGTGAGCGATTTACATGCTGTTCGCGGTACG 1461  
Qy 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429  
Db 1462 GTTTTGTGATGGCCATGACATCAAGGACCTCAATCTCGCTGGCTTCGCCAACAGATC 1521  
Qy 430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLysTyr 449  
Db 1522 TCTTGTGTAGCCAGGAGCCTGTTCTTTTGGCAGCAGCATTTTATAAGAATATTAGGCAC 1581

Qy 450 Gly-----ArgAspAspValThrAspGluValThrAspGluMetGlu 460  
Db 1582 GGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGGAACTCATCTCAG 1641  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1642 AACCGGGCAAAATGGCGAATGCTCATGACTTTATTACTGCTTGGCTGAAGGTTTAG 1701  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1702 ACCAATGTTGGCAGCGTGGCTTTCTCTTTCAGGTGGCCAGAAACAGCGCATTCGATC 1761  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1762 GCCCGTGGCTGTTAGTGACCAAAATCTGCTCTCGATGAAGCTACTTCGGCGCTTG 1821  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1822 GACCAAAATCCGAAGCGGTGTTCAAGCAGCTTTGGAGGGCGAGCTGAAGCCGAACT 1881  
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1882 ACTATTGTGATCGCTCATCGCTTTCCAGCATCAAAACCGCGCACACATTGTGGTTCG 1941  
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1942 GTCAATGGCAAAATGCTGAACAAGAACTCATGATGAATTGGTTGACCGCGAGCGCT 2001  
Qy 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600  
Db 2002 TATCGCAAACTTGTGGAGGCTCAACGATCAATGAACAAGAAAGAGCTGACGCTTGGAG 2061  
Qy 601 ThrTyrSerThrGluArgLysThrAsnSer-----LeuProLeuHisSerValLysSer 618  
Db 2062 GAGCGCAGCTGAGGATCTCAGAAATGCGAGATATGTGCCAAATCAAAATCGGTCAAGC 2121  
Qy 619 IleLysSerAspPhe-----IleAspLys-----AlaGluGluSer 630  
Db 2122 GCATCATCGCATCTCGACGGAAAAACCCACACCAATGTCGCGACGGGACCCCAAGTCT 2181  
Qy 631 ThrGlnSerLysGluIleSer-----LeuProGluValSerLeuLeu 644  
Db 2182 GTTTCAGCGCGATTTCTTTCTAAAGACGCCCGCAAAACATCCGAAATCACTCATTTGG 2241  
Qy 645 LysIleLeuLys-----LeuAsnLysProGluTrpProPheValValLeuGly 660  
Db 2242 AGCTGCTCAAAATTTGTTGCTTCTCAACCGCTTGAATCCCGTACATGTCATCGGT 2301  
Qy 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680  
Db 2302 CTTGTCTTCTCAGTGTAGCTGGTGGTGGCCAAACCCACCAAGCAGTGTATATGCTAAA 2361  
Qy 681 IleIleThrMetPheGly-----AsnAsnAspLysThrThrLeuLysHisAspAlaGlu 698  
Db 2362 GCCATCAGCACACTCTCGCTCCAGAAATCACAATATAGCAAGCTTCGACATGATCGGAT 2421  
Qy 699 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 718  
Db 2422 TTTGTCATATGATGTTCTCGGTGGTGTATCATTTAGTTTATCAGTCATCAGCAGTCAACCAAT 2481  
Qy 719 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 738  
Db 2482 GGTGCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGGAGAGACGCTGCTTT 2541  
Qy 739 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 758  
Db 2542 CGGACGATACTCGTCAAGACATTTGCTTTTTCAGAGAAAGAGAAATAGCACCGCGCT 2601  
Qy 759 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 778  
Db 2602 CTGACCTCTTCTTCTCCACCGAGCAAGCATCTCTCCGGTGTAGCGGTGTGACTCTA 2661

```
QY 779 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 798
Db 2662 GGCACGATCTTGATGACCTCCAGCAGCTAGGAGCGCTATCATTTATGCGCTGGCGATT 2721
QY 799 GlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 818
Db 2722 GGGTGGAAATGGCTTAGTTGTATCTCGGTTGTGCGGTTCTCTCGCATCGGTTTC 2781
QY 819 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 838
Db 2782 TACGATTCATATGCTAGGCCCGATTTCAATCACGCTCCAGCTTGCTATGAGGATCT 2841
QY 839 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 858
Db 2842 GCAAACTTTCCTGCGAGGCTACATCGTCTATCCGACAGTTGCGTCATTAACCCGGAA 2901
QY 859 LysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLys 878
Db 2902 AGGATGCTCGGAGATTACCATGCCAGCTTGACGCACAAGCAGCAGCACGCTTAATC 2961
QY 879 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 898
Db 2962 TCTGCTCTGAGGTCATCCCTGTTATATGCTGCTGCGCAGCACTGTGTTCTTCTCGCTT 3021
QY 899 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 918
Db 3022 CGCTCGGGTTTGGTAGGAGGACACTTCTTGGTCACCAGCATGATGACATTTCCGC 3081
QY 919 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 938
Db 3082 TTCTTTGTTTCTCCGAGATTCTCTTGGTGCTCAATCCGCGGCGCACCGCTTTTCC 3141
QY 939 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 958
Db 3142 TTGTCACACGACATGGCAAGCGCGAAGAAATGCCGCGCCCGAAATCCGACACTGTTGCAC 3201
QY 959 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 978
Db 3202 CGAAGGCCACAAATTGATTAATCTGCTGAGAGGCGCGAGAGCTCGAAGCGGTGGAAGT 3261
QY 979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
Db 3262 GAAATCGAATTTAGGAACGTGCACCTTCAGATACCCGACCGCCCGCAGAACAGCCTGCTCG 3321
QY 999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
Db 3322 CGCGGCTTGGACCTGACCTGAGCCGTGGACAATATGTGCGCTTGTGCGACCCGCGGT 3381
QY 1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
Db 3382 TGTGGCAAGAGTACACCATTTGCATTGCGTTCGAGCGCTTTACGATGCGATGCGGGTCC 3441
QY 1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
Db 3442 ATCCTTGTGTGGAAGGACATAAGTAATAATAATCAACTCTACCAGCAGCTTCTG 3501
QY 1059 AlaIleValProGlnGluProValPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
Db 3502 TCACCTGTGACCCAGGAGCCGACACTGTACCAGGGCCACCATCAAGGAAAAACATCTTACTT 3561
QY 1079 GlyAspAsnSerArgValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1098
Db 3562 GGTATTGTCGAAGATGAGTACCGAAGAAATCTTGATTAGGCTTGCAAGGACGCTAAT 3621
QY 1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
Db 3622 ATCTACGACTTCATCATGCTCCCGAGGGCTTAATACAGTTGTTGGCAGCAGGGA 3681
QY 1119 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
Db 3682 GGCATGTGTGTCGCGGCCAAAGCAACGTGGCCATGTCGCGACCCCTTCTTCGGGAT 3741
QY 1139 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerSerGluLysVal 1158
```

```
Db 3742 CCCAAAAATCCTTTCTTCGATGAAGCGAGCTCAGCCCTCGACTCCGAGTCAGAAAAGGTC 3801
QY 1159 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1178
Db 3802 GTCCAGCGCGCTTTGATGCCCTGCCGAGCCGCAACCAATCCCGCTTGCACACCGA 3861
QY 1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGlu 1198
Db 3862 CTCAGCACGATTCAAAAGCGGACGTTATCTATGTTTTCACCAAGGCAAGATCGTCGAA 3921
QY 1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db 3922 AGCGGAACGCACAGCGAACTGGTCCAGAAAAAGGCCGCGTACTACGAGCTGGTCAACTTG 3981
QY 1219 GlnSerVal 1221
Db 3982 CAGAGCTTG 3990
```

Search completed: March 31, 2003, 15:32:18  
Job time : 455.956 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 8405.14 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-5

Perfect score: 6159

Sequence: 1 MITGLIASLVNGACILPLMLP.....QELLNRDIYKLVNAQSVQ 1222

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cpn2.1/USPTO.spool/US9873409/runat\_27032003.115420.19240/app\_query.fasta\_1.7544  
-DB=EST -QPMF=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US9873409 @CCN 1.1 30544 @runat 27032003.115420.19240 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_nam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	17.7	669	13	BM013981	BM013981 603639619
2	1051	17.1	1033	14	BM926413	BM926413 AGENCOURT
3	1008	16.4	943	9	AL520322	AL520322 AL520322
4	995	16.2	760	13	BM016204	BM016204 603842659
5	918	14.9	545	9	AL040762	AL040762 DKF2P43AC
6	903	14.7	929	14	BQ882401	BQ882401 AGENCOURT
7	856.5	13.9	948	13	BM471690	BM471690 AGENCOURT
8	834	13.5	998	14	BM904842	BM904842 AGENCOURT
9	833.5	13.5	2676	11	AK014319	AK014319 Mus muscu
10	777.1	12.5	894	12	BF584668	BF584668 602098406
11	770.5	12.5	1019	12	BG248052	BG248052 602359987
12	753	12.2	726	12	BG293345	BG293345 602390738
13	752.5	12.2	871	17	AZ682350	AZ682350 ENTKB16TF
14	751	12.2	944	12	BF796582	BF796582 602258463
15	744.5	12.1	872	11	AK020318	AK020318 Mus muscu
16	743.5	12.1	886	17	BH133685	BH133685 ENTNA47TR
17	734.5	11.9	932	17	AZ670821	AZ670821 ENTJUN69TF
18	728.5	11.8	947	17	AZ683753	AZ683753 ENTIL96TF
19	726.5	11.8	886	17	AZ540627	AZ540627 ENTBQ18TF
20	723.5	11.7	880	17	AZ687805	AZ687805 ENTLU52TF
21	721.5	11.7	939	14	BQ720763	BQ720763 AGENCOURT
22	720.5	11.7	897	17	AZ541090	AZ541090 ENTUS67TR
23	719	11.7	780	12	BG587938	BG587938 EST489713
24	711.5	11.6	913	17	BH155700	BH155700 ENTRO54TR
25	709.5	11.5	834	17	AZ548312	AZ548312 ENTRO54TR
26	708.5	11.5	853	17	AZ679807	AZ679807 ENTHI67TR
27	705.5	11.5	823	17	AZ532602	AZ532602 ENTCR50TR
28	697.5	11.3	861	14	BQ717101	BQ717101 AGENCOURT
29	692.5	11.2	1341	11	AY108285	AY108285 Zea mays
30	690.5	11.2	899	17	BH154857	BH154857 ENTRO47TF
31	690	11.2	835	13	BJ349604	BJ349604 BJ349604
32	689.5	11.2	891	17	AZ682250	AZ682250 ENTKT68TR
33	688	11.2	1123	11	AY108485	AY108485 Zea mays
34	687	11.2	715	14	BQ869512	BQ869512 QGD6G09.Y
35	684	11.1	820	13	BJ356540	BJ356540 BJ356540
36	683	11.1	600	13	BM486593	BM486593 pgm2n.pk0
37	682.5	11.1	919	17	AZ690701	AZ690701 ENTVI68TR
38	680	11.0	926	14	BQ123477	BQ123477 EST609053
39	680	11.0	1813	11	U66688	U66688 Homo sapien
40	677.5	11.0	795	12	BF133560	BF133560 601900192
41	674.5	11.0	815	13	BI409057	BI409057 602961127
42	670	10.9	852	17	BH720383	BH720383 BOH2294TF
43	669	10.9	547	13	BJ030711	BJ030711 BJ030711
44	667	10.8	937	17	BH150760	BH150760 ENTQ642TR
45	666	10.8	792	13	BJ352404	BJ352404 BJ352404

# ALIGNMENTS

RESULT 1  
BM013981  
LOCUS  
DEFINITION BM013981 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5415556 5',  
669 bp mRNA linear EST 30-OCT-2001  
mRNA sequence.  
ACCESSION BM013981  
VERSION BM013981.1 GI:16528335  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 669)  
AUTHORS NIH-MGC http://mgc.mci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



```

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
|
|
|
Db 300 AGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAAAGAACTGTGGAA 359
|
|
|
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
|
|
|
Db 360 TTTAAATAATGTTCTTTCATATCCATCAAGACCATCTATCAAGATTCGAAAGGCTCG 419
|
|
|
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
|
|
|
Db 420 AATCTCAGATTAAGTCTGGAGAGACAGTCGCTTGGTCGCTCTCAATGGCATGGGAAG 479
|
|
|
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
|
|
|
Db 480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGCTTTATCATGGTG 539
|
|
|
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
|
|
|
Db 540 GATGAGAATGACATCAGAGCTTTTAAATGTGCGCATATATCGAGACCATATTTGGAGTGGTT 599
|
|
|
QY 433 SerGluGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
|
|
|
Db 600 AGTCAAGAGCCTGTTTGTTCGGGACCATCATCAGTAACATATCAAGTATGACGAGAT 659
|
|
|
QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
|
|
|
Db 660 GATGTGACTGATCAAGAGATGGAGAGCAGCAAGGGAAGCAATGCGTATGATTATC 719
|
|
|
QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
|
|
|
Db 720 ATGGAGTTTCTTAATAA-TTTAATACATTGCTAGGAGGAANAAGAGCTCCAATCAGTGG 778
|
|
|
QY 493 GlyGlnGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
|
|
|
Db 779 AGCCGAACCG-AGGATCCCAATGCTCTGCTTGTAGTTCGAAACCCAGGAT-CTGAAT 836
|
|
|
QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
|
|
|
Db 837 TTATAGTGAAGCTAGCTGCTGCTCGCTTGCAGTTCAGAAACCAAGTCACTGCTC 884
|
|
|
RESULT 3
AL520322 943 bp mRNA linear EST 13-FEB-2001
LOCUS
DEFINITION AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB06YC15 5
prime, mRNA sequence.
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB06YC15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH108"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and

```

cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t  
ORIGIN

Alignment Scores: 2.36e-99 Length: 943  
Pred. No.: 1008.00 Matches: 194  
Score: 80.40% Conservative: 48  
Percent Similarity: 64.45% Mismatches: 58  
Best Local Similarity: 16.37% Indels: 1  
Query Match: 9 Gaps: 0  
DB:

US-09-873-409-5 (1-1222) x AL520322 (1-943)

```

QY 921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
|
|
|
Db 5 TTATATTTCAGCTGTTGCTTGTGTCATGGCGTGGGCAAGTCAGTTTCATTGCT 64
|
|
|
QY 941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
|
|
|
Db 65 CCTGACTATGCCAAAGCCAAATATATCAGCAGGCCACATCATCATCATTTGAAAAAAC 124
|
|
|
QY 961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
|
|
|
Db 125 CCTTTCATTGACAGCTACAGCAGCGAAGGCCTAATGCCGAACACATTTGGAAGAAATGC 184
|
|
|
QY 981 GluPheArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGly 1000
|
|
|
Db 185 ACATTTGGTGAAGTTGTATTCAACTATCCACCAGCCGACATCCAGTGTCTCAGGA 244
|
|
|
QY 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerGlyCysGly 1020
|
|
|
Db 245 CTGAGCTGGAGGTGAAGAGGCGCAGAGCTGGCTCTGTTGGGCGAGCAGTGTGCTGGG 304
|
|
|
QY 1021 LysSerThrSerValGlnLeuGluArgLeuTyrAspProValGlnGlyGlnValLeu 1040
|
|
|
Db 305 AAGAGCACAGTGTGCTGAGCTCTCTGGAGCGGTTCTACGCCCTTGGCAGGGAAGTGTG 364
|
|
|
QY 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIle 1060
|
|
|
Db 365 CTGTATGGCAAGAAATAAAGCAGCTGAATGTTCACTGGCTCCGAGCACACCTGGGCATC 424
|
|
|
QY 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
|
|
|
Db 425 GTGTCCAGGAGCCCATCTGTTTGACTGCGCATTTGCTGAGAACATTTGCTATGGAGAC 484
|
|
|
QY 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
|
|
|
Db 485 AACAGCCGGTGTGTGTACAGGAAGAGATCGTGAGGCGCAGCAAGAGGCGCAACATACAT 544
|
|
|
QY 1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
|
|
|
Db 545 GCCTTCATGAGTCACTGCTCTAATAATATAGCACTAAGTAGAGACAAAGAACTCAG 604
|
|
|
QY 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
|
|
|
Db 605 CTCTCTGGTGGCCAGAAACACACGCAATTCATAGCTCGTCCCTTTGTAGACAGCCTCAT 664
|
|
|
QY 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
|
|
|
Db 665 ATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGATACACAAAAGTGAAGAGTTGTCAA 724
|
|
|
QY 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
|
|
|
Db 725 GAAGCCCTGGCAAAAGCCAGAGAAGCCGCCCTGATTTGTTGCTACCCGCTGTC 784
|
|
|
QY 1181 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGln 1200
|
|
|

```

```

|||||
Db 785 ACCATCCAGAAATGAGAGTCTAATAGTGTGTTCAGAAATGGCAGAGTCAAGGAGCATGG 844
|||||
Qy 1200 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaGlnSe 1220
|||||
Db 845 CAGCATCAGAGCTGTGGCAGAAAGGCACTATTTTCAATGTCAGTCCAGGC 904
|||||
Qy 1220 r 1220
Db 905 T 905

RESULT 4
BM016204 760 bp mRNA linear EST 30-OCT-2001
LOCUS 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
DEFINITION mRNA sequence.
ACCESSION BM016204.1 GI:16530558
VERSION BM016204.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 760)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12068 row: e column: 16
High quality sequence stop: 740.
Location/Qualifiers
1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 4.27e-98 Length: 760
Score: 995.00 Matches: 203
Percent Similarity: 98.07% Conservat: 0
Best Local Similarity: 98.07% Mismatches: 2
Query Match: 16.16% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-5 (1-1222) x BM016204 (1-760)

Qy 330 LysLysProSerIleAspAsnPhSerThrAlaGlyTyrLysProGluSerIleGluGly 349
Db 3 AAGAAACCCAGTATAGTAACCTTTCCACAGCTGGATATAAACCCTGAATCCATAGAAGGA 62
Qy 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
Db 63 ACTGTGAATTTAAAGATGTTCTTCAATTATCATCATCAGACCATCTATCAAGATTCG 122
Qy 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389

```

```

|||||
Db 123 AAAGGCTGAATCTCAAAATTAAGTCTGGAGAGACAGTGCCTTGGTGGTCTCAATGGC 182
|||||
Qy 390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPhe 409
|||||
Db 183 AGTGGGAAGAGTAGCTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTT 242
|||||
Qy 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
|||||
Db 243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGGCGCATTCAGAGACCATATN 302
|||||
Qy 430 GlyValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyr 449
|||||
Db 303 GGAGTGGTGTAGTCAAGAGCGCTGTTTGTTCGGACACCATCATAGTAACAATATCAAGTAT 362
|||||
Qy 450 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 469
|||||
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAGGAGCAAGCAATCGGTAT 422
|||||
Qy 470 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 489
|||||
Db 423 GATTTTATCATGAGTTCCTAATAAATTTAATACATTTGGTAGGGGAAAAAGAGCTCAA 482
|||||
Qy 490 MetSerGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLys 509
|||||
Db 483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTTGGCTCGTGCCTTAGTTTCGAAACCCCAA 542
|||||
Qy 509 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 529
|||||
Db 543 GATTCGTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTGAGAAGCAAGTCAGCTGTTTC 602
|||||
Qy 529 lnAlaAlaLeuGluLys 534
|||||
Db 603 AAGCTGCACCTGGAGAAG 619

RESULT 5
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKFZp434C1815-5', mRNA sequence.
ACCESSION AL040762.1 GI:5409708
VERSION AL040762.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 545)
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Bloecker, et al.)
Unpublished (1999)
Contact: Bloecker H
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
s1 sequence also available.
This clone (DKFZp434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
FEATURES
source

```



```

/lab_host="DH10B"
/Note="Vector: pSPort1; Site 1: NotI; Site 2: SalI"
BASE COUNT 171 a 89 c 138 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 6,63e-90 Length: 545
Score: 918.00 Matches: 181
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.91% Indels: 0
DB: 9 Gaps: 0

US-09-873-409-5 (1-1222) x AL040762 (1-545)

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
DB 3 AGTATAGATAACTTTTCCACAGCTGGATATTAACCTGATCCATAGAGAACTGTGGAA 62
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
DB 63 TTTAAAAATGTTCTTCAATATCCATCAAGACCATCTATCAAGATTCGAAAGGCTCTG 122
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
DB 123 AATCTCAGAAATTAAGTCTCGAGAGACAGTCGCTTGGTCTCAATGCGAGTGGGAAG 182
QY 393 SerThrValValGlnLeuLeuGlnAisGLeuTyrAspProAspAspGlyPheIleMetVal 412
DB 183 AGTACGGTAGTCAGCTTTCGACAGAGTTATATATGATCCGGATGATGGCTTATCATGGTG 242
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
DB 243 GATGAGAATGACATCAGAGCTTTTAAATGTGCGGCATTTATCGAGACCATATTGAGTGGTT 302
QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
DB 303 AGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 362
QY 453 AspValThrAspGluGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPheIle 472
DB 363 GATGTGACTGATGAAGATGGAGAGACGACGACGAGGAAGCAATGCGTATGATTTTATC 422
QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly 492
DB 423 ATGGAGTTTCTTAATAATTAATACATTGTTAGGTAGGGGAAAAAGAGCTCAATGAGTGA 482
QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
DB 483 GGGCAGAAACAGAGGATCGCAATTGCTCGTCTGCTTAGTTTCGAAACCCCAAGATTCTGATT 542
QY 513 Leu 513
DB 543 TTA 545

RESULT 6
BO882401
LOCUS
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
5', mRNA sequence.
ACCESSION BO882401
VERSION BO882401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

```

```

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.
Location/Qualifiers
1. 929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 7,64e-88 Length: 929
Score: 903.00 Matches: 188
Percent Similarity: 73.02% Conservative: 61
Best Local Similarity: 55.13% Mismatches: 58
Query Match: 14.66% Indels: 34
DB: 14 Gaps: 3

US-09-873-409-5 (1-1222) x BO882401 (1-929)

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThr-Val11 195
DB 1 GCGTATGCAAAAGCTGGAGCAGTACTGGAAGAGGCTCTGGCAGCAATTAGAACTGGTAT 60
QY 195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 215
DB 61 TGCATTGGAGGACAAAGAAAGAACTTGAA----- 91
QY 215 rAlaTyrPheTyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrTh 235
DB 92 -----AGGTACAA 99
QY 235 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 255
DB 100 CAAAATTTAGAAAGCTTAAAGAAATTGGGATAAAGAGCTATTACAGCCAATATTTC 159
QY 255 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 275
DB 160 TATAGTGTGCTTCTCTCTCATCTATGCTATGCTATGCTCTGCGCTTCTGGTATGGGAC 219
QY 275 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 295
DB 220 CACCTTGGTCTCTCAGGGAA-----TATTCTATGGACAAAGTACTCACTGTATCTT 273
QY 295 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 315
DB 274 TTTCTGATTAAATGGGGCTTTTAGTGTGGACAGGAGCTCTCCAAGCATTTGAACATTTGC 333
QY 315 alLeaAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 335
DB 334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTGATAAATAAGCAAGTATTGA 393
QY 335 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 355
DB 394 CAGCTATTGCAAGAGTGGGCACAAACCAAGATAATAATTAAGGGAATTTTGGAAATTCAGAA 453

```

```

QY 355 nValSerPheAsnTyrProSerArgProSerIleLeuLysGlyLeuAsnLeuAr 375
Db 454 TGTTCACTTCAGTTACCACTCGAAGAAGATTAGATCTTGAAGGGTCTGAACCTGAA 513
QY 375 gileLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 395
Db 514 GGTGCAGATGGGCAGACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGACACAAC 573
QY 395 lValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 415
Db 574 AGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGGACA 633
QY 415 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 435
Db 634 GGATATTAGACCAATAATTAAGTTCTTACGGGAATCATTTGGTGTGTGAGTCAGGA 693
QY 435 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValTh 455
Db 694 ACCTGTATTGTTGCCACCAAGATAGTGAACCAATTCCTATGGCCGTGAAATGTCCAC 753
QY 455 rAspGluLeuMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetClup 475
Db 754 CATGGATGAGATTGAGAAGCTGTCAGGAAGCAATGCCCTATGACTTTATCATGAAC 813
QY 475 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly---G 494
Db 814 TGCTCATATAATTGACACCTGTTGGANAAGAGGGGCCAATTGAGTGTGGGCGAC 873
QY 494 lnLysGlnArgIleAlaAlaAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 511
Db 874 AAACCAAGGATCTCCCTGGCAGCTGGCCCTGGTTCCCAACCCCAAGATCCTC 928

RESULT 7
BM471690 948 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
DEFINITION 5', mRNA sequence.
ACCESSION BM471690
VERSION BM471690.1 GI:18520732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 948)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12233 row: b column: 14
High quality sequence stop: 569.
FEATURES
Location/Qualifiers
1..948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5539117"
/clone_lib="NIH MGC 72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 285 a 187 c 194 g 282 t

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 1e-82 Length: 948  
 Score: 856.50 Matches: 186  
 Percent Similarity: 89.67% Conservative: 5  
 Best Local Similarity: 87.32% Mismatches: 14  
 Query Match: 13.91% Indels: 8  
 Dbs: 13 Gaps: 1

US-09-873-409-5 (1-1222) x BM471690 (1-948)

```

QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
Db 205 GTTTCTTTTAGTGAATCCATAGCAGTATTGTCATTGGAGCAGCAGTCCCTCACTTTGAA 264
QY 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
Db 265 ACCTTCGCAATAGCCCGAGGAGCTGCTTTTCATATTTCAGGTTATTGATAAGAACCC 324
QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAACCTGAATCCATAGAGAACTGTGGA 384
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
Db 385 TTTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGCTG 444
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
Db 445 AATCTCAGAAATTAAGTCTGGAGACAGTCGCTTGGTCTCAATGGCAGTGGGAAG 504
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
Db 505 AGTACGGTAGTCAGCTTCTGCAGAGTTATATGATCCGGATGATGGCTTATCATGGTG 564
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 432
Db 565 GATGAGAATGACATCAGAGCTTTTAAATGTGGGCAATTATCGAGACCATATTTGGAGTGT 624
QY 432 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 451
Db 625 TAGTCAAGAGCCCTGGTTTGTTCGGGACCATCATCAGTAACAATATCAAGTATGGGACG 684
QY 451 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 470
Db 685 AGATGATGGAGCTGATGACAAATGGAAGAACCAACAGGGAAACCAAGGGAATGAA 744
QY 471 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 489
Db 745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAGGGGAAAAAAGGACCTTCA 804
QY 489 nMetSerGlyGlyGlnLysGlnArg 497
Db 805 AGGAATGGAAGGCGGCAAAACCCCAAG 829

RESULT 8
BM904842 998 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
DEFINITION 5', mRNA sequence.
ACCESSION BM904842
VERSION BM904842.1 GI:19355221
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

```

Qy	745	AspIlealaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla	764
Db	61	GATGTGAGTTGGTTGATGACCCCTAAAAACACCACTGGAGCATTCAGCTACCGAGCTCGCG	120
Qy	765	IleaspIlealaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn	784
Db	121	AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTCTGTAATATACCCAGAAT	180
Qy	785	AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe	804
Db	181	ATAGCAAAATCTTGGGACAGAAATATATATCTTCATCTATGTTGGCAACTAAACACTG	240
Qy	805	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	824
Db	241	TTACTCTTAGCAATGTACCCATCATCTAGTGCATAGCAGGAGTCTTGAAATGAAATGTTG	300
Qy	825	ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu	844
Db	301	TCTGGACAAGCACTGAAAGATAGAAAAGAACTAGAGGTCCTGGGAAGATCGCTACTGAA	360
Qy	845	AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet	864
Db	361	GCAATAGAAAACCTCCCAACCGTTGTTCTTTGACTCAGGACAGGAAGTTTGAACATATG	420
Qy	865	TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly	884
Db	421	TATGCTCAGAGTTTGCAGGTACCCATACAGAAACTCTTTGAGGAAGACACATCTTTTGA	480
Qy	885	SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe	904
Db	481	ATTACATTTTCCTTACCCAGGCAATGATATATTTTTCCTATGCTGGATGTTTCCGGTTT	540
Qy	905	GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr	924
Db	541	GGAGCCCTACTTGGTGGCACATAAACTCATGAGCTTTTGAGGATGTTCTGTTAGTATTTTCA	600
Qy	925	AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer	944



897 CTTACCTGATGCTGCTGTCACACCGCCCTCATGGAGTGGGACCCCTGATGGGC 956  
 164 ArgMetValIleSerLeuSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183  
 957 TCAGGCTCCGAAGCTGCTCCAGTGTCCAGGACAGATTCCAGGGCCACAGGTGA 1016  
 184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluysGlu 203  
 1017 GCAGATGAGGCCCTTGCCCAATGTTCCGACTGTGCGGCCTTCCGATGAGGAAGAGGAG 1076  
 204 LeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyr 223  
 1077 GAGGAACGC-----TATCAAGCAGAA 1097  
 224 LeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaVal 243  
 1098 CTGAGCTCATGCTGCTGT-----AAAGCAGAA 1127  
 244 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe----- 261  
 1128 CTGGGC-----AGGGCATCGCTTGTTCAGAGGCTCTCCAAACATCGCTTCAACTGT 1181  
 262 PheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuLeuAsnGly 281  
 1182 ATGGTCTTGGGCACC-----CTGTTTATGGGGCTCCCTTGTGGCTGGACAG 1229  
 282 Glu---ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 1230 CAGCTCAAGAGGGGACCTCATGCTCTCTGCTGGTCCAGACAGTACAGAGTCT 1289  
 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly---- 319  
 1290 ATGGCCAGCTCTCTCTCTGCTGTTTGTGTGAG-----GTGGTACGTGGGCTG 1334  
 320 ---AlaAlaPheHisIlePheGlnValIleAspLysIleProSerIleAspAsnPheSer 338  
 1335 AGTGGGGGACCGAGCTCTTCCGAATACATGGCCCTGAGCCCTGCTCATCCCA-----TTG 1388  
 339 ThrAlaGlyTyr-----LysProGluSerIleGluGlyThrValGluPheLysAsnVal 356  
 1389 ACCGGGGCTACGTGCATCCCAACAGGACATTCGTGGTTCATACCTTCCAAATGTC 1448  
 357 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 376  
 1449 ACCTTCAGTACCCCTGCAGACCTGCTCAATGCTCAAGGACTTCACCTGAAGCTG 1508  
 377 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 396  
 1509 CCCTCTGGCAAGATTGCGCTCTTGTGGCCAGTCTGGGGGAGGAAGACCAAGTTGCC 1568  
 397 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 416  
 1569 TCCTCTGGAACGCTCTATGACCTTGAAGTGGCTGGCTGGCTGGATGGGATGAC 1628  
 417 IleArgAlaLeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGlu 435  
 1629 CTGCGAATCTCAACCCCTCTGCTCCGGGCCAGGTCTATGTTTTCATCACCAGGAG 1688  
 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr 455  
 1689 CCAGTCTGTTTGCACCAACCATCATGGGAATATCCGATTGGGAAGCTGGATGCTTC 1748  
 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475  
 1749 GATGAAGAGGTGTACACAGCTGCACGAGAGCAATGCCACAGGTTTCATCAGACGCTTC 1808  
 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495  
 1809 CCGATGGCTACAGCACTGTGTTGTGGAGGGGGCACACCTTGTCTGGTGGCCAGAAG 1868  
 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuLeuAspGlu 515

Db 1869 CAGCGCTAGCCATCGCAGCTGCCCTCATCAAGACGCCACAGTCTGATCTCTGGACGAG 1928  
 Qy 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAla 535  
 Db 1929 GCCACCAAGTGTAGTGTGACAGTGTGTTACAGAGGCGCTTGGACCGGCGC 1988  
 Qy 536 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 555  
 Db 1989 AGTGTGTCGCGCAGCCGTGTGTGTCATTCGCCACCGCTTAGTACTGTCTCGTGCAGCCAC 2048  
 Qy 556 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 575  
 Db 2049 TCATCATCTGTATGCCCAATGCCAAGTCTGTGAGCTGGGACCCACCAAGAACTCCTT 2108  
 Qy 576 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 590  
 Db 2109 AAAAAAGGGGGCTCTATTTCAGAGCTTATCCGAGACAAACCTGT 2153  
 RESULT 10  
 BF584668 894 bp mRNA linear EST 12-DEC-2000  
 LOCUS 602098406F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4218385 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF584668 GI:11658386  
 VERSION BF584668.1  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 894)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapps-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Cloning Distribution by: Incyte Genomics, Inc.  
 Cloning Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM9798 row: 1 column: 02  
 High quality sequence stop: 651.  
 Location/Qualifiers  
 FEATURES  
 source 1..894  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4218385"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI CGAP Library."  
 BASE COUNT 263 a 203 c 242 g 185 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,18e-73 Length: 894  
 Score: 771.00 Matches: 171  
 Percent Similarity: 79.72% Conservative: 57  
 Best Local Similarity: 59.79% Mismatches: 48  
 Query Match: 12.52% Indels: 15  
 DB: 12 Gaps: 3  
 US-09-873-409-5 (1-1222) x BF584668 (1-894)  
 Qy 293 ValPhePheSerValIleHisSerTyrCysIleGlyAlaAlaValProHisPheGlu 312  
 Db 5 GTCTCTCTTCTGCTGTTAT-TGGAGCATTCAGTGTGGACAGGCATCTCCAAATATTGAA 63

QY 313 ThrPheAlaIleAlaAArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 332  
 Db 64 GCCTTCGCCAATGACGAGGAGGAGCTTATGAAGTCTTCAAAATAATGATAAAGCCC 123  
 QY 333 SerIleAspAsnPheSerThrAlaGlyTyrProGluSerIleGluGlyThrValGlu 352  
 Db 124 AGTATACAGAGCTTCTCAAGAGTGGGCAAAACACGACACATACAGAAATCTGGAA 183  
 QY 353 PheIleAsnValSerPheAsnTyrProSerArgProSerIleIleValIleLeuLysGlyLeu 372  
 Db 184 TTTAAGAATATTCATCTCAGTTACCCATCTCGAAAGAGTTCAGATCTTGAAGGGCCCTC 243  
 QY 373 AsnLeuArgIleLysSerGlyThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392  
 Db 244 AATCTGAAGTGAAGAGCGACAGCGTGGCTGGTGGCAACAGTGGCTGGGAAA 303  
 QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412  
 Db 304 AGCACAACTGTCCAGCTGATGCAAGGCTCTAGCACCCCTAGATGGCATGGTCAGTATC 363  
 QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432  
 Db 364 GACGGACAGACATCAGAACCAATGTGAGGTATCTGAGGGAGATCATGTGTGGTG 423  
 QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452  
 Db 424 AGTCAGAACTGTGCTGTTTGGCCACCATGCGCGAGAACATTCGCTATGGCCGAGAA 483  
 QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472  
 Db 484 GATGTCAACCATGATGATGATGAGAAAGCTGTCAAGGAGCCCAATGCTATGACTTCATC 543  
 QY 473 MetCluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492  
 Db 544 ATGAACCTGCCCCACCAATTTGACACCTGTGTGGTGAAGAGGGCGCAGCTGAGTGGG 603  
 QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512  
 Db 604 GGACAGAAACAGAGAAATCGC-ATTGCCCGGGCCCTGTGTCGCAATCCCAAGTGCCTTC- 661  
 QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla-Le 532  
 Db 662 GTGACCAAGCCACTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTGAGGGCGCACCT 720  
 QY 532 uGluLysAlaSerLysGlyArgThrIle-----ValValAlaHisArgLe 548  
 Db 721 AGATAGGGTACAAAAGGCGNGGACCAATTTGGGAGAGCTCATCGCTGTTCGTCGAAGCT 780  
 QY 548 uSerThrIleArgSerAlaAspLeuIleVal-----ThrLeuLysAspGlyMetLeuAl 566  
 Db 781 GACGGCCATCGGGCCGAGGGGGG-CTTGTGGGAGACACTCTGNAAGC-----TC 830  
 QY 566 aGluLysGlyAlaHis 571  
 Db 831 TGAAAAAGGTACTCAT 846  
 RESULT 11  
 BG248052  
 LOCUS 602359987F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG248052  
 VERSION BG248052.1 GI:12757867  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1019)  
 NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgbbs-r@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10334 row: n column: 21  
 High quality sequence stop: 650.  
 FEATURES  
 source  
 1..1019  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strains="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4488404"  
 /clone\_lib="NCI\_CGAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 276 a 262 c 283 g 196 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,13e-73 Length: 1019  
 Score: 770.50 Matches: 154  
 Percent Similarity: 79.17% Conservative: 36  
 Best Local Similarity: 64.17% Mismatches: 47  
 Query Match: 12.51% Indels: 3  
 DB: 12 Gaps: 1  
 US-09-873-409-5 (1-1222) x BG248052 (1-1019)  
 QY 977 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 996  
 Db 4 GAAGGAATGTGCAATTTAGTGGAGTCTGTTCAACTATCCACCCGACCCAGCATCCCA 63  
 QY 997 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 1016  
 Db 64 GTCTTTCAGGGGTGAGCCTTGAGGTGAAGAGGGCCAGACGCTGGCCCTGTGGGCGC 123  
 QY 1017 SerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGln 1036  
 Db 124 AGTGGCTGGGGAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183  
 QY 1037 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1056  
 Db 184 GGATCAGTGTTCATGATGGCAAGAAATAAGCACTGATGTCCTCCAGTGGTCCGAGCA 243  
 QY 1057 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1076  
 Db 244 CAGCTGGCATTTGTGTCCCAAGAGCCATTTCTTTGACTGCAGCATCGCAGAGAACATT 303  
 QY 1077 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAsnAla 1096  
 Db 304 GCCTACGGAGACAAACAGCGGGTCTTATGAGGAGATTGTGAGGGGAGCAGCAGGAG 363  
 QY 1097 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1116  
 Db 364 GCCAACATCCACCATTCATCGACTCGTACCTGATATAATACCAACAGAGTAGGAGAC 423  
 QY 1117 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1136  
 Db 424 AAAGGCACCTCAGCTGTGGGTGGGAGAGAGCGCATCGCCATCGCAGCGCCCTCGTC 483  
 QY 1137 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 1156  
 Db 484 AGAGCCCTCACATTTTACTTCTGGAGCAAGCAACATCAGCTCTGGATACAGAAAGTGA 543

Qy	1051	valGlnTrpLeuArgSerGlnleAlaIleValProGlnGluProValValPheAsnCys	1070
Db	122	GTCCAGTGGCTCGGAGCTCAACTGGGCATTTGTGCCAGGAACCATTCCTCTTTTGACTGC	181
Qy	1071	SerIleAalagluasnIleAtyrGlyAspAsnSerArgValValProLeuAspGluIle	1090
Db	182	AGCATCGCAGAGAACAATCCCTATTGGAGACAACACCGGGTCTGTCCTCATGATGAGATT	241
Qy	1091	LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr	1110
Db	242	GTGAGGGCAGCCAAGGAGGCCAACATCCACCCTTCATCGAGAGCGCTCCCCAAAATAAT	301
Qy	1111	AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla	1130
Db	302	AACACAAGAGTAGGAGACAAGGGGACGACGAGCTCTCTGGGGGCCAGAAACAGAGGATTGCC	361
Qy	1131	IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla	1150
Db	362	ATCGCCCGAGCCCTCATACAGACAGCGCTGGGTCTTACTGCTGGATGAAGACCCAGTCAGCT	421
Qy	1151	LeuAspAsnAspSertGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg	1170
Db	422	CTGGATCTGAGAGTGAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGAGGCGCGC	481
Qy	1171	ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValva	1190
Db	482	ACCTGATTGGCAATCGCTCACCGCTGTCCACATCCACAGACGGCGCACTTGATCGTGGT	541
Qy	1190	IleuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA	1210
Db	542	GATTGAGACGCGCAGGTCNAGSAGCAGCGCACCCCACCAGCAGCTGCTGGCGCAGAGG	601
Qy	1210	spleTyrPhe-LysLeuValAsn 1217	
Db	602	GCATCTATTCTCAACTGGTGCAAC 625	
RESULT	13		
ACCESSION	AZ682350	AZ682350	871 bp DNA linear GSS 14-DEC-
LOCUS	ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.		
DEFINITION			
VERSION	AZ682350	AZ682350.1 GI:11819496	
KEYWORDS	GSS.		
SOURCE	Entamoeba histolytica.		
ORGANISM	Entamoeba histolytica		
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.		
AUTHORS	1 (bases 1 to 871)		
TITLE	Lofcus B., Van Aken S. and Fraser, C.		
JOURNAL	Determination of clone end sequences from Entamoeba histolytica		
COMMENT	HMI:IMSS sheared DNA library Unpublished (2000) Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: bjloftus@tigr.org Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library Seq primer: M13-Forward Class: shotgun High quality sequence start: 16 High quality sequence stop: 860. Location/Qualifiers 1. .871 /organism="Entamoeba histolytica" /strain="HMI:IMSS" /db_xref="taxon:5759" /clone_libs="Entamoeba histolytica Sheared DNA" /note="Vector: pHSO1; Site 1: Bst I; Constructed at ThermoFisher Scientific"		
FEATURES			
source			





QY 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 219 CTGAGGTGAAGAAAGCCAGACACTAGCCCTGGTGGGAGCAGTGGCTGTGGGAAGAGC 278  
 QY 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAsp 1042  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 279 ACGGTGGTCCAGCTCTGAGCGGTTCTACGACCCCTTGGCGGGACAGTGGCTTCTCGAT 338  
 QY 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 339 GGTCAAGAAAGCAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398  
 QY 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 399 CAGGAGCCTATCTCTTACTGCTCAGCATTTGCCGAGATATTGCTTATGGAGACAACAGC 458  
 QY 1083 ArgValValProLeuAspGluLeuLysGluAlaAsnAlaAsnIleHisSerPhe 1102  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 459 CGGGTTGTATCATCAGATGAATCTGTGAGTGCAGCCCAAGCTGCCAACATACATCCTTTC 518  
 QY 1103 IleGluGlnLeuProGluLysTyrAsnThrGlnValGlnLeuLysGlyAlaGlnLeuSer 1122  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 519 ATCGAGACCTTACCCACAAATATGAAACAGACTGGGAGATAGGGGACTCAGCTCTCA 578  
 QY 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProlysisIleLe 1142  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 579 GGAGGTCAACAACAGAGGATTGCTATTGCCGAGCCCTCATCAGACAACCTCAATCTCT 638  
 QY 1142 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal-----G 1160  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 639 CCTGTTGGATGAAGCTACATCAGCTCTGATACCTGCACAAAGTGAACACGGTTGTCCCA 698  
 QY 1160 lnhIsAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 1180  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 699 GAAGCCCTGTACACAGACAGAGCGCGACCTGCTTGTGATGTCTCACCCTGTC 758  
 QY 1180 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 1200  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 759 AAC--ATCCAGATGCAGACTT-ATAGTGGTGTGCCAACGGGAGAAGTCAAG----- 808  
 QY 1200 lyThrHis 1202  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 Db 809 --ACACAT 814  
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
 RESULT 15  
 AK020318  
 LOCUS  
 DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.  
 ACCESSION AK020318.1 GI:12860872  
 VERSION  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE 2  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374

11042159

PUBMED  
REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076861

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C. C., Wilmink, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

5 (bases 1 to 872)  
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

TITLE  
JOURNAL

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer

## COMMENT

adapter of sequence [5',  
GAGAGAGAGATCTCGAGTAAATTAATATCCGCCCCCCCCC 3']. cDNA was cleaved  
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after  
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3',  
end: BamHI. Host: DH10B

## FEATURES

## source

## Location/Qualifiers

1..872  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:9230106F14"  
/db\_xref="MGD:MG1:1911565"  
/db\_xref="taxon:10090"  
/clone="9230106F14"  
/sex="male"  
/tissue type="epididymis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"

## CDS

86..>872  
/notes="data source:SPTR, source key:P21447, evidence:ISS  
putative  
similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN  
3) (MDR1A)"  
/codon\_start=1  
/protein\_id="BAB32070.1"  
/db\_xref="GI:12860873"  
/db\_xref="MGD:MG1:1924956"  
/translation="MANERTNGLOETNORYGPOBQVKGNOAVGPIEIPFADNL  
DIVLMTGLIASINGATVPLMSVLGEISDHLINGCLVQTNRTKYQNGSQTCEKLE  
DIIVLYYIGIGAAALIFGYVOISFWITAAQTRIRKQFPHSILAODISWFGSD  
ICELNTRMTGDKLDCGIDGKPLMFQNIISGLVSLIKSLVLSVLSTPLI  
MASSALCSRMIIISLTSEKLDAYSGAVAEALSSIQTVTAFGA"

BASE COUNT 262 a 178 c 187 g 245 t  
ORIGIN

## Alignment Scores:

Pred. No.:	1.69e-70	Length:	872
Score:	744.50	Matches:	151
Percent Similarity:	80.37%	Conservative:	21
Best Local Similarity:	70.56%	Mismatches:	27
Query Match:	12.09%	Indels:	15
DB:	11	Gaps:	2

US-09-873-409-5 (1-1222) x AK020318 (1-872)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	230	ATGACTCTGGGAATATTAGCATTCATGATAATGAGCCCGCTCCTTTAATGTCCTG	289
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	290	GTTTATAGGAGAAATAAGTGATCATTTAATTAAATGATGCTAGTACAAACTAACAGAACT	349
Qy	41	SerPhe-----PheArgLeuThr	46
Db	350	AAATATCAGAACTGTTCTCAGACTCAAGAAAGCTGAATGAAGATATCATTTGTTGACT	409
Qy	47	LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu	66
Db	410	CTATATTATATTGGAATAGGAGAGCTGCCCTCATTTTGGCTAATGACAGATTTCCTTC	469
Qy	67	TrpIleIleThrAlaAlaArgGlnThrIysArgIleArgIysGlnPhePheHisSerVal	86
Db	470	TGGGTCACTACGCGCCGCAACCAACAGAGATCCGAAACAGTTTTTTCATTCAATT	529
Qy	87	LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet	106
Db	530	TTGGCACAAGACATCAGCTGGTTTGTATGGCAGTGACATCTGTGAACCTTAACACCCGATG	589
Qy	107	Thr---AspIleAspIleIleSerAspGlyIleGlyAspIleIleAlaLeuLeuPheGln	125
Db	590	ACTGGTGACATCAACAAACTCTGTGATGGTATGGAGATAAGATCCCTCTGATGTTTCAG	649
Qy	126	AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIysGlyTyrIysLeuThr	145

4 . . .

Db	650	AAACATATCTGGGTTTTCTATTGGCTTGGTATAAGTTTATATAAAAGCTGGAAACTCTCC	709
Qy	146	LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet	165
Db	710	CTGGTGGTTCTGTCTACATCTCTCTCATATAGGCTTCATCGGCACCTGTGCTTAGGATG	769
Qy	166	ValIleSerLeuThrSerIysGluLeuSerAlaTyrSerIysAlaGlyAlaValAlaGlu	185
Db	770	ATTATCTCATTCACCAAGAGAGCTGGATGCCATTATCCAAAGCTGGGGCTGTGGCTGAA	829
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAla	199
Db	830	GAAGCCTTGTCTCAATCAACCAACAGTCACAGCCTTTGGAGCC	871

Search completed: March 31, 2003, 13:50:58  
Job time : 8432.26 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 1194.3 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-5

Perfect score: 6159

Sequence: 1 MITGLIASLVNGACPLMLP.....QSLRLNRDIYFKLVNAQSVQ 1222

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=cpn2\_1/USPTO spool/US09873409/runat\_27032003\_115416\_19216/app\_query.fasta\_1.7544  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CCN 1.1 4247 @RUNAT 27032003 115416 19216 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	5737	93.1	3699	24	ABK83223	Human transporter
2	3533.5	57.4	4279	22	AA03506	Dog P-glycoprotein
3	3533	57.4	4195	22	AAF86128	Cynomolgous monke
4	3530.5	57.3	3860	21	AAZ49332	Human wild-type mu
5	3530.5	57.3	3860	24	ABA94365	Human BCRP DNA rel
6	3530.5	57.3	4279	22	AA03504	Dog P-glycoprotein
7	3530.5	57.3	4279	22	AA03505	Dog P-glycoprotein
8	3525.5	57.2	4317	22	AA03488	Dog P-glycoprotein
9	3525.5	57.2	4317	22	AA03489	Dog P-glycoprotein
10	3524.5	57.2	4186	22	AAF86127	Cynomolgous monke
11	3521.5	57.2	3860	21	AAZ49333	Human G185V mutant
12	3521.5	57.2	3860	24	ABA94366	Human BCRP DNA rel
13	3521.5	57.2	4349	22	AAH57442	Human intestine ce
14	3521.5	57.2	4646	21	AAZ94738	Human ATP binding
15	3521.5	57.2	4646	24	AAD38994	Human mdrl gene.
16	3521.5	57.2	4646	24	ABL68592	Kidney cancer rela
17	3521.5	57.2	4646	24	ABL68880	Kidney cancer rela
18	3521.5	57.2	6505	17	AAAT13394	Hybrid vector p8F-
19	3521.5	57.2	8630	21	AAZ24041	Retroviral M4 mdr-
20	3521.5	57.2	8630	21	AAZ24042	Hybrid vector p8F-
21	3520	57.2	4264	19	AAV66533	Retroviral vector
22	3520	57.2	4264	19	AAV66534	Mutated human P-gl
23	3519.5	57.1	4669	8	AAV70752	Mutated human P-gl
24	3517	57.1	4425	21	AAZ52048	Sequence of human
25	3515.5	57.1	4378	11	AAQ04522	Rat multidrug resi
26	3515	57.1	4369	21	AAZ52047	Multidrug Resistan
27	3514	57.1	3840	24	ABL91687	Rat multidrug resi
28	3514	57.1	3988	21	AAZ88973	Human polynucleoti
29	3512.5	57.0	4669	14	AAQ52726	Human MDR-1 DNA.
30	3510.5	57.0	4646	15	AAQ72872	Sequence of human
31	3503.5	56.9	4669	19	AAV32645	Human multidrug re
32	3503.5	56.9	4669	24	ABK52041	Human P glycoprote
33	3503.5	56.9	4189	21	AAZ49334	cDNA encoding huma
34	3503	56.9	4189	24	ABA94367	Murine multidrug r
35	3503	56.9	4189	24	ABA94367	Mouse BCRP DNA rel
36	3499	56.9	4313	14	AAQ38950	Mouse multidrug re
37	3499	56.8	4788	21	AAZ49335	Murine multidrug r
38	3494	56.7	4233	24	ABA94368	Mouse BCRP DNA rel
39	3494	56.7	4233	22	AAF27498	Rat mdrlb2 (multis
40	3481	56.5	3924	21	AAZ94742	Rat mdrlb2 multidr
41	3481	56.5	3924	21	AAZ88974	Human ATP binding
42	3481	56.5	3924	21	AAZ88974	Human MDR-3 DNA.
43	3453	56.1	4254	24	ABK63517	Gene #2299 used to
44	3405	55.3	3912	24	ABK63653	Rat sequence diffe
45	2844.5	46.2	4776	21	AAZ94744	Rat sequence diffe
						Human ATP binding

ALIGNMENTS

RESULT 1	
ABK83223	
ID	ABK83223 standard; cDNA; 3699 BP.
XX	ABK83223;
AC	ABK83223;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX	
KW	Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW	neurological disorder; muscle disorder; immunological disorder; cancer;
KW	scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW	cell proliferative disorder; cervical cancer; breast cancer;
KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW	myoclonic dystrophy; catatonia; endocrine disorder; diabetes;
KW	Grave's disease; gastrointestinal disorder; Crohn's disease;
KW	renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;  
 KW protozoal infection; helminthic infection; cardiovascular disorder;  
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-NOV-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-252790P.

XX (INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwaia M;

XX Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for  
 PT diagnosing, preventing, and treating disorders associated with an  
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular  
 PT or renal disorders

XX Claim 5; Page 172; 178pp; English.  
 CC The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
 CC a recombinant polynucleotide comprising a promoter sequence operably  
 CC linked to the TRICH polynucleotide, a cell transformed with the  
 CC recombinant polynucleotide, a transgenic organism comprising the  
 CC recombinant polynucleotide, an isolated antibody that binds specifically  
 CC to TRICH, and screening for compounds which bind to TRICH, modulate  
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.  
 CC The polypeptides are useful for diagnosing, treating, and  
 CC preventing transport, neurological, muscle, immunological disorders  
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or  
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
 CC catatonla), endocrine disorders (e.g. diabetes, Grave's disease),  
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
 CC (e.g. Good pasture's syndrome), vital, bacterial, fungal, parasitic,  
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
 CC other diseases and disorders detailed in the specification. They can also  
 CC be used in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of transporters and  
 CC ion channels. TRICH or its fragments may also be used in screening for  
 CC compounds that specifically bind to and modulate the activity of TRICH.  
 CC The polynucleotides can be used to create knock-in humanised animals or  
 CC transgenic animals to model human disease. The present sequence  
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

XX Alignment Scores:

Prd. No.: 0 Length: 3699  
 Score: 5737.00 Matches: 1156  
 Percent Similarity: 93.78% Conservative: 4

Best Local Similarity: 93.45% Mismatches: 9  
 Query Match: 93.15% Indels: 68  
 DB: 24 Gaps: 4

US-09-873-409-5 (1-1222) x ABK83223 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 145 ATGATCCTGGGTATACCTGACATCAGTCTGTTCAATGGAGCCTGCTCTTAAATGCCACTG 204  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
 DB 205 TGTATAGGAGAAATAGTATACCTTATTAGTGGATGTCTAGTCCACACTTAAACAACA 264  
 QY 41 SerPhePheArg-----LeuThr 46  
 DB 265 AATTATCAGAACTGACTCAGTCTCAAGAGAACTCAATGAAGATATGACTCTGTTGACC 324  
 QY 47 LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
 DB 325 CTGTATTATTGTTGGAATAGTGTGCTGCTTGTATTTTGGTTTACATACAGATTCTCTTG 384  
 QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleAargLysGlnPhePheHisSerVal 86  
 DB 385 TGGATTATTAACCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 444  
 QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
 DB 445 TTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 504  
 QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
 DB 505 ACAGATGACATTGACAAAATCAGTGATGTTATGGAGATAAGATTCTCTGTGTTTCAA 564  
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145  
 DB 565 AACATCTCTACTTTTTTCGATTGGCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACCTCAC 624  
 QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165  
 DB 625 CTAGTACTCTATCCACGTCCTCTCTTATTAAGGCTTCAGCGCAGCAGCAGCAGCAGCAG 684  
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
 DB 685 GTCATCTCTATTCACAGTAGTAAGGAATTAAGTGCCTATTCAAAGCTGGGCTGGCAGAA 744  
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln 205  
 DB 745 GAAGTCTTGTTCATCAATCCGAAACAGTCATAGCCTTTTAGGCCCCCAGGAGAAAGAACTCAA 804  
 QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTyrLeuLeu 225  
 DB 804 -----AGGTATACACAGAACTCAAGATCAAAAGGATTTTGGC 843  
 QY 226 SerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245  
 DB 805 -----AGGTATACACAGAACTCAAGATCAAAAGGATTTTGGC 843  
 QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265  
 DB 844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTTGGTGGCTGTGTACTTCTTTATGAATGA 903  
 QY 266 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285  
 DB 904 ACCTATGGACTTGTCTTTTGGTATGAACCTCTCTGATTCTTATGAGAACCTGGATAT 963  
 QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305  
 DB 964 ACCATCGGACTGTCTTCTGCTGTTTCTTTAGTGTAACTCATAGTAGTATTATTCATTGGA 1023  
 QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325  
 DB 1024 GCAGCAGTCCCTCCTTGAACCTTTCGCAATAGCCCGAGGAGCTGCTTTCATATTTTC 1083

Qy	326	GlnValIleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGlu	345
Db	1084	CAGGTATTGATAAGAAACCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAA	1143
Qy	346	SerIleGluGlyThrValGluPheLysAenValSerPheAenTyrProSerArgProSer	365
Db	1144	TCCATAGAAGAACTCTGGAAATTTAAAAATGTTTCTTTCAATATTATCCATCAAGGCCATCT	1203
Qy	366	IleLysIleLeuLysGlyLeuAenLeuAglIleLysSerGlyGluThrValAlaLeuVal	385
Db	1204	ATCAAGATTCTGAAGGTCGAATCTCGAAATTAAAGTCTGGAGAGACAGTCGCCTTGGTC	1263
Qy	386	GlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	405
Db	1264	GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGCTTATATGATCCG	1323
Qy	406	AspAspGlyPheIleMetValAspGluAenAspIleAArgAlaLeuAenValArgHisTyr	425
Db	1324	GATGATGGCTTTATCATCTGGTGGATGAGAAATGCATCAGAGCTTTAAATGTGCGGCATTAT	1383
Qy	426	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAen	445
Db	1384	CGAGACCATAATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGAGCACCACCATCAGTAAC	1443
Qy	446	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu	465
Db	1444	AAATATCAAGTATCGACGAGATGATGTGACTGATGAAGAGATGGAGAGACGACGAGGGA	1503
Qy	466	AlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyGlu	485
Db	1504	GCAAAATCGGTATGATTTATCATGAGGTTCCTAAATAAATTTAATACATGTTGATGGGGAA	1563
Qy	486	LysGlyAlaGlnMetSerGlyGlnLysGlnAglIleAlaIleAlaArgAlaLeuVal	505
Db	1564	AAAGGAGCTCAATGATGGAGGCGAGAAACAGAGAGATCGCAATTTGCTGTGGCTTAGTT	1623
Qy	506	ArgAenProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	525
Db	1624	CGAAACCCCAAGATTCTGATTTAGATGAGGCTACGCTCTGCCCTGGATTTCAGAAAGCAAG	1683
Qy	526	SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla	545
Db	1684	TCAGCTGTTCAAGCTGCACCTGGAGAAGCGCAGCAAAAGTCCGACTACAATCGTGATAGCA	1743
Qy	546	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	565
Db	1744	CACCGACTTTCTACTATTCGAAGTCGAGATTTGATTGTGACCCCTAAAGGATGGAAATCGCT	1803
Qy	566	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	585
Db	1804	CGCGAANAAGAGCGCATGCTGAACTTAATGGCAAAACGAGGTCTATATATTTCATCTGTG	1863
Qy	586	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	605
Db	1864	ATGTCACAGGATTTAAAAAAGCTGATGAACAGATGGAGTCAATGCATATTCTACTGAA	1923
Qy	606	ArgLysThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAsp	625
Db	1924	AGAAAGACCAACTCACTTCTCTGCACCTCTGTGAAGAGCATCAAGTCAAGCTTCATTGAC	1983
Qy	626	LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys	645
Db	1984	AAGGCTCAGGAATCCACCCAAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAA	2043
Qy	646	IleLeuLysLeuAenLysProGluTyrProPheValValLeuGlyThrLeuAlaSerVal	665
Db	2044	ATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGACATTTGGCTCTGTT	2103
Qy	666	LeuAenGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe	685
Db	2104	CTAAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTTGCAAAAATTTATAACCATGTTT	2163

QY	686	GlyAsnAsnAspIysThrThrLeuIysHisAspAlaGluIleTyrSerMetIlePheVal	705
DB	2164	GGAAATATGATAAAAACCATTTAAAGCATGATGAGAAATTTATTTCATGATATTGCTG	2223
QY	706	IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla	725
DB	2224	ATTTTGGGTGTATTGTCTTTGTCACTTATTTCATG	2259
QY	726	GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp	745
DB	2260	-----CAGGAT	2265
QY	746	IleAlaTrpPheAspGluIysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle	765
DB	2266	ATTGCCCTGGTTGATGAAAGGAAAAACAGCAGAGAGCGCTTGACAAACAATATTAGCCATA	2325
QY	766	AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla	785
DB	2326	GATATGACACAAATTCAGGAGCACAAGGTTCCAGATTGGCGTCTTAAACACAAANAATGCA	2385
QY	786	ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu	805
DB	2386	ACTAACATGGACTTCACGTATCATTTCTTTATATGATGGAGATGACATCTTCGT	2445
QY	806	IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr	825
DB	2446	ATTTCTGAGTATTGCTCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACT	2505
QY	826	GlyPheAlaAsnIysAspIysGlnGluLeuLysHisAlaGlyValIleAlaThrGluAla	845
DB	2506	GGATTGCCAACAAAGATAAGCAAGAACTTTAAGCATGCTGGAAAGATAGCAACTGAAGCT	2565
QY	846	LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr	865
DB	2566	TGGAGATATACGTACTAGTGTCAITTAACAGGGAAAAAGCCCTTCGAGCAATGTAT	2625
QY	866	GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer	885
DB	2626	GAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGACACAGATTATTGGGAAGC	2685
QY	886	CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly	905
DB	2686	TGTTATGCATTTCAGCCATGCGCTTTATATATTTCCTATGCGGGCAGGGTTTCGATTTCGA	2745
QY	906	AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla	925
DB	2746	GCCTATTTAATTCAGCTGAGCAATGACCCACAGAGGGCATGTTTCATAGTTTTTACTGCA	2805
QY	926	IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys	945
DB	2806	ATTGCATATGAGCTATGGCCATCGGAGAAACCGCTGTTTGGCTCCTGAAATATTCCAAA	2865
QY	946	AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer	965
DB	2866	GCCAAATCGGGGGCTGGCCATCTGTTTGCTTGTGTGAAAGAAACCCAAAATATAGACAGC	2925
QY	966	ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal	985
DB	2926	CGCAGTCAGAGAGGAAAAAGCCAGACACATGTGAAGGGAAATTAAGATTTCGAGAGATC	2985
QY	986	SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle	1005
DB	2986	TCCTTTCTTATCCATGTCGCCAGATGTTTTTCATCTCCGTGGCTTATTCCTCAGTATT	3045
QY	1006	GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal	1025
DB	3046	GAGCGAGGAAAGACATGATGATTTGTGGGGAGCAGCGCTGTGGGAAAGACACTCTCTGT	3105
QY	1026	GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValIlePheAspGlyValAsp	1045
DB	3106	CAACTTCGCAGAGACTTTATGACCCCGTCGAGGAGCAAGTCTCTTTGATGGTGTGAT	3165
QY	1046	AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro	1065

```
Db 3166 GCAAAAGAAATTGAATGTACAGTGGCTCGGTTCCCAATAGCAATCGTTCCTCAAGAGCCT 3225
Qy 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
Db 3226 GTGCTCTTCAACTGCACATGTGTGAGAACATCGCTATGTCACACACCCGCTGGTG 3285
Qy 1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
Db 3286 CCATTAGATGAGATCAAGAAGCGCAATGCGCAATATCCATCTTTATTATGAGGT 3345
Qy 1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGln 1125
Db 3346 CTCCTCAGAAATACACACACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGCGGCCAG 3405
Qy 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1145
Db 3406 AAACAAAGACTAGCTATGCAAGGGCTCTTCTCCAAAACCCAAAATTTATTGTTGGAT 3465
Qy 1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
Db 3466 GAGCCACTTCAGCCCTCGATATATGACAGTGAGAGGTGTTGAGCATGCCCTTGATAAA 3525
Qy 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
Db 3526 GCCAGGACGGAGGACATGCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCA 3585
Qy 1186 AspLeuIleValValLeuHisAsnGlyValIleLysGluGlnGlyThrHisGlnGluLeu 1205
Db 3586 GATTGTGATAGTGGTTCTGCAATGGAAGATTAAGGAACAGCACTCATCAAGAGCTC 3645
Qy 1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
Db 3646 CTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696

RESULT 2
AAD03506
ID AAD03506 standard; cDNA; 4279 BP.
XX
AC AAD03506;
XX
DT 13-JUN-2001 (first entry)
DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
XX MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype D) protein"
FT allele replace (91, T)
FT /*tag= b
FT allele replace (607, C)
FT /*tag= c
FT allele replace (1001, T)
FT /*tag= c
FT allele replace (3459, A)
FT /*tag= c
XX
PN WO200123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
```

```
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
DR P-PSDB; AAE00310.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 9; Page 102-107; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDRI) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX
SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4279
Score: 3533.50 Matches: 683
Percent Similarity: 73.13% Conservative: 237
Best Local Similarity: 54.29% Mismatches: 269
Query Match: 57.37% Indels: 69
DB: 22 Gaps: 9

US-09-873-409-5 (1-1222) x AAD03506 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAAATGCTGCCATCATCCATGGAGTGCACCTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTTCAAGAAACAAACCTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CAGTTATATTAATTAATGAAAGTATTACGAACATATACAAATTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCACCGATGCTTATTATTACAGTGGATCGGTCTGCGTGTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTCATTCTGGTGGTGGCAGGAGGAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATGCGACAGGAGATTTGGCTGGTTCACGTGATGACGTT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAATCAATGAGGAATGGGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAAATTGGAATGTTCTTCAATCAATGAGCAACATTTTTCACCGGTTTATATAGTGGGTTT 646
```



QY 139 ValYleGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuLeuMetAlaSer 158  
Db 647 ACACGGTGGTGAAGCTAACCCCTGTGATTTGGCCATCAGCCCTGTCTCTGGACATTCA 706  
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 707 GCGCCATCTGGCCAAAGATACTATCTTCTTACTGATAAAGAACTCTTGGCCTATGCA 766  
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 767 AAAGCTGGAGCAGTACTGAGGAAGCTTACGAGCAATCAGAACTGTGATTCCTTTGGA 826  
QY 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218  
Db 827 GGCAAAAGAAAGAACTTGAA----- 847  
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeu 238  
Db 848 -----AGGTACAAACAAAAATTTA 865  
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
Db 866 GAAGAAGCTTAAGCAATTTGGGATTAAGAAAGCTATCAGGCCCAACATTTCTATTGTGCC 925  
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278  
Db 926 GCCTTCTTATGATCATGCATCATATGCTCTGCCTTCTGGTATGGACCTCTCTGGTC 985  
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
Db 986 CTCTCCAGTGAA-----TATACTATTGGCAAGTACTCACTCTCTCTTTCTGTATT 1039  
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
Db 1040 ATGGGGCTTTTATGATTGGACAGGCATCCCAAGCAATTGAAGCATTTGCAACCGCAAGA 1099  
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
Db 1100 GGAGCAGCTTATGAATCTTCAGAGTAATTTGCAATAAACCAAGCATTGACAGCTATTGC 1159  
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358  
Db 1160 AAGAGTGGACATAAACCAGATAATATTAAAGGAAATTTGGAATTTCAAAAATGTTCACTTC 1219  
QY 359 AsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSer 378  
Db 1220 AGTTACCCCTTCTCGAAAGAAAGTTAAGATCTTAAAGGGTCTCAACCTGAAAGGTTGAGT 1279  
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
Db 1280 GGGCAGACAGCTGGCGCTGGTTGGGACAGTGGCTCGGGGAAGACGACCGGTGCAGCTG 1339  
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
Db 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTGCTGTTATGATGACAGGACATTAGG 1399  
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
Db 1400 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGGTGTGTGAGTACGAGGCTGTGTTG 1459  
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458  
Db 1460 TTTGCCACACGATAGCTGAAACATTCGCTATGGCCGCGGAAATGTACACCATGGATGAG 1519  
QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
Db 1520 ATTGAGAAAGCTGTTAAGGAAGCAATGCCTATGATTTTATCATGAACCTACCTAATAAA 1579  
QY 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
Db 1580 TTTGACACTCTGGTTGGAGAGAGGGGCCCGCAGCTGAGTGGTGACAGAAACAGAGATC 1639

QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
Db 1640 GCCATTGCTCGGCCCTGTTCCCAACCCCAAGATTCTTCTGCTGATGAGGCAACGTCA 1699  
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538  
Db 1700 GCTCTGGACACTGAAGTGAAGCAGTGGTTCAGTGGCCCTGATGAAGGCCAGAAAGGC 1759  
QY 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558  
Db 1760 CGGACTACCATTTGTGATAGCTCATCGTTTGTCTACAGTTCTGTAATCCCGATGTCATTGCT 1819  
QY 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578  
Db 1820 GGTTTTGTATGGATCATTTGTGGAGAAAGGAAATCATCATGAACTCATGAAGAAGAG 1879  
QY 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598  
Db 1880 GGCATTACTTCAAACTTGTCAATGCAG--ACMAGAGGAAATGAAATTTGATTAGTAA 1936  
QY 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
Db 1937 AATGCCACTGGTGAATCCAAAGTGAAGTGCCTTGGAAATG---TCTCCAAAAGAT 1993  
QY 619 IleLysSerAspPheIle----- 624  
Db 1994 TCAGGGTCCAGTTTAAATAAAAGAAAGATCAATCGCAGGAGTATACATGCACCACAAAGC 2053  
QY 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643  
Db 2054 CAAGCAGAAAGCTGGTACAAAGAGAGCCTGAATGAGAAATGTACTCCAGTTTCCCTTC 2113  
QY 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663  
Db 2114 TGGAGATCTGAGCTGAATCACTGAATGGCCTATTATTGCTGGTGGTATTTTGT 2173  
QY 664 SerValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThr 683  
Db 2174 GCTATTATAACCGGAGCTGCACACCAGCATTTTCAATAATATTTTCAAGGATTATAGG 2233  
QY 684 MetPhe---GlyAsnAsnAspLysThrLeuLysPheHisAspAlaGluIleTyrSerMet 702  
Db 2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACACAGCAATAGTAACTGTTTCTGTA 2293  
QY 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722  
Db 2294 TTGTTTCTAGTCTCTGGAATATTCTTTTATTACATTTTCTCCAGGCTTCACATTT 2353  
QY 723 GlyArgAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742  
Db 2354 GGCAAGCTGGGAGATCCTCAAGCGCTTCGATACATGGTTTTTTCAGATCCATGCTG 2413  
QY 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762  
Db 2414 AGACAGAGTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2473  
QY 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782  
Db 2474 CTTGCCAATGATCGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCTTGTCTGCTATACC 2533  
QY 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802  
Db 2534 CAGAAATAGCAAACTCTGGGACAGGCAATTTATATCTTAACTATGTTGTTGGCAATTA 2593  
QY 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822  
Db 2594 ACATTTTACTCTTAGCAATTTGACCATTTGCAATAGCAGGAGTTGTTGAATGAAA 2653  
QY 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842  
Db 2654 ATGTTCTGGCAACAGCTCTGAAAGATAGAAAGAGCTGAGAGAGCTGGGAAGATTGCT 2713  
QY 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862



US-09-873-409-5 (1-1222) x AAF86128 (1-4195)

Qy	1	MetIleuGlyIleIleuAlaSerLeuValAsnGlyAlaCysLeuProIleuMetProLeu	20
Db	250	ATGGTGGTGGAACTTTGGCTGCATCATCCATGGAGCTGGACTTCTCTCATGATGCTG	309
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys	33
Db	310	GTGTTGGAGCATGACGGATACCTTTGCCAATTCAGGNAATTTAGGAGATTTAGGAGCT	369
Qy	34	LeuValGlnThrAsnThrTyrSer-----	41
Db	370	CTGTTGTTTAAACAACACTAATAGCAGTAATATCACTGATACAGTCCCGTCATGAATCTG	429
Qy	42	-----PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaIleuIle	58
Db	430	GAGGAAGATATGACACAGGTATGCTATTATTATACAGTGGAAATTTGGTGGGGTCTGGT	489
Qy	59	PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIle	78
Db	490	CTGCTTACATTCAGGTTTCATTTTGCTGCTCGGCAGCTGGAGACAATATACACAAAT	549
Qy	79	ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp	98
Db	550	AGAAAACAGTTTTTTCATGCTATAATGCACAGGAGATAGCGTGGTTGATGTGCACGAT	609
Qy	99	IleGlyGlnLeuAsnThrArgMetThr----AspIleAspLysIleSerAspGlyIleGly	117
Db	610	GTTGGGGAGCTTAAACACCCGGCTTACAGATGATGTCCTCCAAATTAATGAAGAAATTTGGT	669
Qy	118	AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly	137
Db	670	GACAAAATGGAATGTTCTTTCACTCAATGGCAACATTTTCACTGGGTTTATAGTAGGA	729
Qy	138	LeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuMetAla	157
Db	730	TTTACACGTGTTGGAAGCTAAACCTTGTGATTTTGGCCATCAGTCTGTTCTTGGACGTG	789
Qy	158	SerAlaAlaLysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr	177
Db	790	TCAGCTCAGTCTGGGCAAGATACCTGCTTCATTTACTGATTAAGAACTCTTACGTTAT	849
Qy	178	SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe	197
Db	850	GCAAAAGCTGACGAGTAGCTGAAGAGGTCTTGGCAGCAATTTAGAACCTGTGATTCATTT	909
Qy	198	ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyr	217
Db	910	GGAGGACAAAAGAAAGAACTCGAA-----	933
Qy	218	PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn	237
Db	934	-----AGGTACAACAAATAAT	948
Qy	238	LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly	257
Db	949	TTAGAAGAAGCTAAAGAATTGGGATTAAGAAGACTATTACAGCCAAATTTTCTATAGGT	1008
Qy	258	AlaValTyrPhePheMetAsnGlyTyrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeu	277
Db	1009	GCTGCTTCTGCTTATCTATGATCTTATGCTCTGCGCTCTGCTATGGGACCACTTG	1068
Qy	278	IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal	297
Db	1069	GTCTCTTCAAGGAA-----TATTCTATGGCAAGTACTCACGTGATTTCTTTCTGTA	1122
Qy	298	IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla	317
Db	1123	TTAATTGGGCTTTAGTGTGTGACAGGCAATCTCCAGCAATTGAAGCAATTGCAAATGCA	1182
Qy	318	ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe	337

Db	1183	AGAGGACGACGCTTTTGGAAATCTCTTCAAGATAAATTGATAATAAGCCAAGTATTGTGACAGCTAT	124
Qy	338	SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValgluPheLysAsnValSer	357
Db	1243	TCGAAGAGTGGGCACAAACACAGATAATAATTAAGGGAAATTTGGAATTCAGAAATGTTCCAC	1302
Qy	358	PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys	377
Db	1303	TTCAgTtTACCcATCTCGAAAGAAGTTTAAGATCTTTGAAGGGCTCGAACTCGAAGGTGCAG	1362
Qy	378	SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln	397
Db	1363	AGTGGGCAGACGGTGGCCCTGGTTGGAACACACGGCTGTGGGAAGAGACAAACGGTCCG	1422
Qy	398	LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle	417
Db	1423	CTGATGCAGAGGCTTTATGACCCACAGAGGCATGCTCAGTGGTTGATGGCAGAGATATT	1482
Qy	418	ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal	437
Db	1483	AGGACCAATAACCTAAGGTTTCTACCGGGAATCATCGGTGTGTGAGTCAGGAACCTGTA	1542
Qy	438	LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu	457
Db	1543	TTGTTTCCACCACGATAGCTGAAACATTCGCTATGTCGTGAAGATGTCCACCATGAT	1602
Qy	458	GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn	477
Db	1603	GAGATTGGAAGAAGCTGTCGAAGGAAGCAATGCTATGACTTTATCATGAACTGCCTCAG	1662
Qy	478	LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg	497
Db	1663	AAATTTGCACCCCTGGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGGCAGAAAGCAGAGG	1722
Qy	498	IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr	517
Db	1723	ATCGCCATTGACGTGCCCTGGTTTCGAACCCCAAGATCCTCCTGCTGGACGAGGCCACG	1782
Qy	518	SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys	537
Db	1783	TCAGCCTTGCACACAGAAAGTGAAGCAGTGGTTTCAGGTGGCTCTGGATAAGGCCAGAAA	1842
Qy	538	GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeulle	557
Db	1843	GGTCGGACCACTTGTGATAGCTCATCTGTTGTTCTACGGTTTCGTAATGCCACGTCATC	1902
Qy	558	ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlalys	577
Db	1903	GCTGGTTTCGATGATGAGTCAATTGTGGAGAAGGAAATCATGATGAGCTCATGAAAGAG	1962
Qy	578	ArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle-----	590
Db	1963	AAAGCATTTACTTCAAACTTGTCACAAATGCAACAGCAGGAGAAATGAATGAATTAGAA	2022
Qy	591	LysLysAlaAspGlu-----GlnMetGluSerMetThrTyrSerThrGluArgLys	607
Db	2023	AATCAGCTGATGAATCCAAAAGTGAAATTGATACCTCTCGAAATGCTCTCACATGATTCA	2082
Qy	608	ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAsp	622
Db	2083	GGATCCAGTCTAATAAGAAAAGATCCACTCGTAGGAGTGTCCGTGGATCAACAGGCCAA	2142
Qy	623	PheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSer	642
Db	2143	-----GACGAAGAAGCTTAGTACCAAGAGGCTCTGGATGAAGAAGTATACCTCCAGTTTCC	2196
Qy	643	LeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeu	662
Db	2197	TTTTGGAGGATTATGAAGCTAAATTAACTGAGTGGCCTTATTTTGTGTTGTTGTTGTTAT	2256
Qy	663	AlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIle	682
Db	2257	TGTGCCATTAATAATGAGGTCTGCACCAACAGATTTTCAGATTAATATTTTCAAGAATTATA	2316

QY 683 ThrMetPhe---GlyAenAenAspLysThrThrLeuLysHisaspAlaGluIleTyrSer 701  
 Db 2317 GGGATTTTACAGAATATGATGCGCAACAAACAGACAGATAGTACTGTTTCA 2376  
 QY 702 MetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPhe 721  
 Db 2377 CTATTGTTCTAGTCCCTGGAATGTTCTTTATTACATATTTTCTTCAGGCGTTCA 2436  
 QY 722 TyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisIleuAlaPheLysAlaMet 741  
 Db 2437 TTTGGCAAGCTGGAGAGATCCTCACCAAGCGGCTCGATACATAGTGTTCGATCCATG 2496  
 QY 742 LeuTyrGlnAspIleAlaTyrPheAspGlyLysGluAenSerThrGlyGlyLeuThrThr 761  
 Db 2497 CTCAGACAGATGTGAGCTGGTGTGATGACCCCTAAACACACCACTGGAGCAITGACTACC 2556  
 QY 762 IleuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeu 781  
 Db 2557 AGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGCTATAAT 2616  
 QY 782 ThrGlnAenAlaThrAenMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGlu 801  
 Db 2617 ACCAGAAATATAGCAAACTCTGGGACAGGAATATATATCTTAATCTATGTTGGCAA 2676  
 QY 802 MetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThr 821  
 Db 2677 CTGACACTGTACTCTTAGCAATGTGTACCATCATCTGCAATAGCAGAGTGTGAAATG 2736  
 QY 822 AlaAlaMetThrGlyPheAlaAenLysAspGlnGlyLeuLysHisAlaGlyLysIle 841  
 Db 2737 AAAATGTTGCTGGCAACACACTGAAAGATAAGAAAGAACTAGAAAGGTGCTGGAGATC 2796  
 QY 842 AlaThrGluAlaLeuGluAenIleArgThrIleValSerLeuThrArgGlyLysAlaPhe 861  
 Db 2797 GCTACTGAAGCAATAGAAAACCTTCGAACTGTTGTTCTTGACTCAGGAGCAGAAGTTT 2856  
 QY 862 GluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGln 881  
 Db 2857 GAACATATGATGATCAGATGTTGCGAGTACCACATACAGAACTCTTTGAGGAAGACAC 2916  
 QY 882 IleIleGlySerCysTyrAlaIlePheSerHisAlaPheIleTyrPheAlaTyrAlaGly 901  
 Db 2917 ATCTTTGGAATCAGTGTTCCTTCACGAGCAATGATGATTTTCTTCTATGCTGATGT 2976  
 QY 902 PheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIle 921  
 Db 2977 TTCGGGTTTGGAGCCTACTTGGTGGCAATAGTCTCATGAGCTTGGAGATGTTCTGTGA 3036  
 QY 922 ValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaPro 941  
 Db 3037 GTATTTTCACTGTGTCTTGTGTCATGCGCGTGGGCAAGTCATGTTTCAATTTGCTCT 3096  
 QY 942 GluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysPro 961  
 Db 3097 GACTATGCCAAAGCAAAGTATCAGCAGCCACATCATCATCATGATGATGATAAAGCCCT 3156  
 QY 962 AsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGlyLysAsnLeuGlu 981  
 Db 3157 TTGATTACACAGCTACACACAGAGGCGCTAAAGCCGCAACATTTGGAAGAAATGTCA 3216  
 QY 982 PheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeu 1001  
 Db 3217 TTTAATGAAGTTGTTATCACTATCCACCGACTGGACATCCAGTGTTCAGGGGCTG 3276  
 QY 1002 SerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLys 1021  
 Db 3277 AGCTGGAAGTGAAGAGGCCAGACGCTGGCCCTGCTGGGACAGCTGCTGTGGGAAG 3336  
 QY 1022 SerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPhe 1041  
 Db 3337 AGCACGGTGTCCAGCTCCTGGAGCGGTCTTATGACCCCTTGGCGGGAAAGTCTGCT 3396

QY 1042 AspGlyValAspAlaLysGluLeuAenValGlnTrpLeuArgSerGlnIleAlaIleVal 1061  
 Db 3397 GACGGCAAAAGAAATAAGCAACTGAATGTTTCACTGCTCGAGCACACCTGGGCACTGTTG 3456  
 QY 1062 ProGlnGluProValLeuPheAenCysSerIleAlaGluAenIleAlaTyrGlyAspAsn 1081  
 Db 3457 TCCAGAGGCCCATCTCTGTTGACTGTCAGCATTTAGTGAACAATTCCTATGAGACAAC 3516  
 QY 1082 SerArgValValProLeuAspGluIleLysGluAlaAaenAlaAlaIleHisSer 1101  
 Db 3517 AGCCGGTGTGTCTCAGGAAGAGATCGTAGGCGCAAGAGGACCAATATACAGCC 3576  
 QY 1102 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValIcilyLeuLysGlyAlaGlnLeu 1121  
 Db 3577 TTCATCGAGTCACTGCTCTTAATAATATAGCACCAGAGTAGGAGACAAGAACTCAGCTC 3636  
 QY 1122 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaIleLeuLysProLysIle 1141  
 Db 3637 TCTGGTGGCCAGAAACAGCATTCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3696  
 QY 1142 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 1161  
 Db 3697 TTGCTTTTGATGAACCAATCAGCTCTGATGATGATGATGATGATGATGATGATGATGAT 3756  
 QY 1162 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 1181  
 Db 3757 GCCTCGCAAAAGCCAGAGAGCGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3816  
 QY 1182 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 1201  
 Db 3817 ATCCAGAATGCAGACTTAATAGTGTGTTTTCAGAAATGGCAGAGTCAAGGAGCACGCG 3876  
 QY 1202 HisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
 Db 3877 CATCAGAGCTGCTGGCAGAAAGGCATCTATTTTCAATGCTGCTGCTGCTGCTGCTGCT 3933

RESULT 4  
 AAZ49332 standard; cDNA; 3860 BP.  
 ID AAZ49332 standard; cDNA; 3860 BP.  
 XX AAZ49332;  
 AC AAZ49332;  
 DT 14-MAR-2000 (first entry)  
 XX Human wild-type multidrug resistance-1 (MDR-1) cDNA.  
 DE Multidrug resistance; MDR-1; P-glycoprotein;  
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassaemia;  
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
 KW cytokine; wild-type; ds.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FX 1..3843  
 FT CDS /\*tag= a  
 FT /product= "Human wild-type MDR-1 protein"  
 FT mutation replace (553..555, GTT)  
 FT /\*tag= b  
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given  
 in AAZ49333"  
 XX WO9961589-A2.  
 XX 02-DEC-1999.  
 PD 27-MAY-1999; 99WO-US11825.  
 PF 28-MAY-1998; 98US-0086988.  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.







QY 1187 LeuileValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206  
 Db 3724 TTAATAGTGGTGTTCAGAAATGCGAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTG 3783

QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
 Db 3784 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 5  
 ABA94365  
 ID ABA94365 standard; DNA; 3860 BP.  
 XX AC ABA94365;  
 XX DT 26-MAR-2002 (first entry)  
 XX DE Human BCRP DNA related seq Id No. 1.  
 XX KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;  
 KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;  
 KW cardiant; gene therapy; ds.  
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT 1..3843  
 FT CDS /\*tag= a

XX PN WO200192877-A2.  
 XX PD 06-DEC-2001.  
 XX 30-MAY-2001; 2001WO-US17459.  
 XX 31-MAY-2000; 2000US-0584586.  
 XX 29-MAY-2001; 2001US-0866866.  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX PI Sorrentino B, Schuetz J;  
 XX WPI; 2002-114368/15.  
 XX P-PSDB; ABB07266.

XX PT Identifying a stem cell, for treating e.g., muscular dystrophy,  
 PT myocardial infarction, Parkinson's disease, or neurodegenerative  
 PT disorders, comprises detecting the expression of an ATP transport  
 PT protein (BCRP) by a cell  
 XX Disclosure; Page 53-55; 87pp; English.

XX CC The invention provides a method of identifying and/or isolating a stem  
 CC cell that involves detecting the expression of an ATP transport protein  
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample  
 CC comprising stem cells. The isolated stem cells may be used in the  
 CC treatment of diseases such as muscular dystrophy, degenerative liver  
 CC disorder, myocardial infarction, Parkinson's disease, degenerative  
 CC disorders of the brain, and for tissue regeneration or replacement.  
 CC Haematopoietic cells can be used in bone marrow transplants (e.g., for  
 CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
 CC also be used as cell targets in gene therapy protocols. The present  
 CC sequence represents a sequence related to the BCRP for which no relevant  
 CC information has been provided in the specification.

XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:  
 Pred. No.: 0 Length: 3860  
 Score: 3530.50 Matches: 682  
 Percent Similarity: 74.00% Conservative: 246  
 Best Local Similarity: 54.39% Mismatches: 263  
 Query Match: 57.32% Indels: 63

DB: 24 Gaps: 10  
 US-09-873-409-5 (1-1222) x ABA94365 (1-3860)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 151 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210

QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 211 GTGTTTGGAGAAATGACAGATATCTTTCAAATGCAGGAAATTTAGAACATCTGATGTCA 270

QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 271 AACATCAGCTAATAGAAATGATATCAATGATACAGGGTCTTTCATGAATCTGGAGGAGAC 330

QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 331 ATGACAGATATGCTTATTTACAGTGGAAATTTGGTGGGGTGTGTTGCTGCTTAC 390

QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 391 ATTCAGGTTTCATTTTGGTGGCTGGAGACAAATACACAAATATTAGAAACAG 450

QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
 Db 451 TTTTTCATGCTATAATGCGACAGAGATAGGCTGTTTGTATGTGCACGATGTTGGGGAG 510

QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 511 CTTAACACCCGACTTACAGATGATGTCCTTAAGATTAAATGAAGGTATTGTTGACAAAATT 570

QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 571 GGAATGTTCTTTTCAGTCAATGCGAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 630

QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 631 GGTTCGAAGCTAACCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 691 GTCTGGGCAAGATACATCTCTTCATTTCTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 751 GGAGCAGTAGCTGAGAGAGCTTTGGCAGCAATTAGAACTGTGATTCATTGGAGACAA 810

QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
 Db 811 AAGAAAGAACTTGAA----- 825

QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
 Db 826 -----AGGTACAAACAAATAATTAGAAAGAA 849

QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
 Db 850 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTC 909

QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
 Db 910 CTGCTGATCTATGATCATCTTATGCTCTGGCTTCTGTTATGGGACCACTTGGTCTCTCA 969

QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 Db 970 GGGGAA-----TATTCTATTGGACAAGTACTCAGTGTATCTTTCTGTATTAAATGGG 1023

QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
 Db 1024 GCTTTTAGTGTGTGGCAGGCAATCTCCAGCATTTGAAGCATTTGCAATGCAAGGAGCA 1083

QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340



Db 1084 GCTTATGAAATCTTCAAGATAAATGATAATAAGCCAGTAGTATTGACAGCTATTTCGAAGAGT 1143  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1144 GGGCACAACACAGATATATTAAGGGAATTTGGAATTCAGAAATGTTCTACTTCAGTTAC 1203  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1204 CCATCTCGAAGAGTATAGATCTTGAAGGCTGAACTGGAAGTGCAGAGTGGGCAG 1263  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1264 ACGTGCCCTGTTGGAACACAGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATGCAG 1323  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1324 AGGCTCTATACCCACAGAGGGGATCGTCAAGTGTGATGCAGCAGGATATAGGACCAT 1383  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1384 AATGTAAGTTTCTACGGGAATCATTTGGTGTGGTGAAGTCAAGNACCTGATTGTTGCC 1443  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
Db 1444 ACCAGATAGCTGAAACATTCGTATGGCGGTGAAAAATGTCACCATGGATGAGATTGAG 1503  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1504 AAGCTCTCAAGGAAGCAATGCTATGACTATATCATGAAGTGCCTCATAAATTGAC 1563  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1564 ACCCTGTTGGAGAGAGGGGCCAGTTGAGTGTGGCAGAGCAGAGGATCGCCATT 1623  
QY 501 AlaArgAlaLeuValArgAsnProLysIleIleLeuLeuAspGluAlaThrSerAlaLeu 520  
Db 1624 GCACGTGCCCTGTTGCAACCCCAAGATCTCTGCTGGATGAGGCCACGTGAGCCCTTG 1683  
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuLysAlaSerLysGlyArgThr 540  
Db 1684 GACACAAAGCAGGAGTGGTTTCAGCTGGCTGTGATAGGCCAGAAAGGTGCGACC 1743  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1744 ACCATTGTGATAGCTCATGTTGTTGTACAGTTCGTATGCTGACGTCACTGCTGGTTTC 1803  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACATCATGAAAGAGAAAGGCATT 1863  
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
Db 1864 TACTTCAAACTGTGCACATGCACAGCAGGAAATGAAGTTGAATAGAAATGCGAGCT 1923  
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 1924 GATGAATCCAAAGTGAATGTATGCTTGGAAATGCTTCAATGATTCAAGATCCAGT 1983  
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
Db 1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAG 2043  
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeu 647  
Db 2044 CTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTTCTTTTGGAGGATTATG 2103  
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
Db 2104 AAGCTAAATTAACCTGAATGGCCCTATTTGTTGGTGTATTTTGTGCCATTATAAT 2163  
QY 668 GlyThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn 687

Db 2164 GGAGGGCTCGAACCCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223  
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706  
Db 2224 AITGATGATCCTCGAACAACACGACAGATAGTAACCTGTTTCTACTATTGTTCTAGCC 2283  
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726  
Db 2284 CTTGGAATTAATTTCTTTTATTACATTTTCTCAAGTTTTCACATTTGCAAGCTGGA 2343  
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746  
Db 2344 GAGATCCTCACCAGCGCTCCGATACATGTTTCCGATCCATCTCAGACAGGATG 2403  
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766  
Db 2404 AGTTGTTTGTATGACCTAAACACCATCGAGCATTCAGTACCAGGCTCGCAATGAT 2463  
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786  
Db 2464 GCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTTGTCTAATATTACCCAGAAATAGCA 2523  
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806  
Db 2524 AATCTTGGGACAGGAATAATATATCTTCTATCTATGTTGGCACTAAACACTGTTACTC 2583  
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826  
Db 2584 TTAGCAATTTGATCCCATCATTCGCAATAGCAGGAGTTGTTGAATGAATGTTGCTGGA 2643  
QY 827 PheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846  
Db 2644 CRAAGCACTGAAAGATAGAAAGAACTAGAAAGTGTCTGGGAAGATCGTACTGAAGCAATA 2703  
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866  
Db 2704 GAAATCTTCGAACCTGTTTCTTCTGACTCAGGAGCAAGTTTGAACATATGTTATGCT 2763  
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886  
Db 2764 CAGAGTTGCGAGGTACCATACAGAAACTCTTTGAGAAAGCACACATCTTTGGAATTACA 2823  
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906  
Db 2824 TTTTCTTCAACCCAGCAATGATGATTTTCTCTATGCTGGATGTTTCCGTTTGGAGCC 2883  
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926  
Db 2884 TACTTGTGGCACATAAACTCATGAGCTTTGAGGATGTTTCTGTTAGTATTTTTCAGCTGTT 2943  
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946  
Db 2944 GTCTTTGGTGCCTAGCCGTGGGGCAAGTCAGTTCATTGCTCTGACTATGCCAAGGCC 3003  
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProHenileAspSerArg 966  
Db 3004 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC 3063  
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986  
Db 3064 AGACGCGAGGCGCTAATGCGGAACACATTTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123  
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006  
Db 3124 TTCAACTATCCACCGACCGGACATCCAGTCTCTCAGGGAAGTGGAGGTTGAGG 3183  
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026  
Db 3184 AAGGGCCAGACGCTGCTGCTGGTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG 3243  
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAla 1046  
Db 3244 CTCCTGGAGCGGTTCTACGACCCCTTTGGCAGGGAAAGTGTCTGTTGATGCAAGAAATA 3303





```
Db 2774 TACATGATGACAGAGTTTGCAGTACCATACAGAACTCTTTGAGGAAGACACATC 2833
QY 883 ILeGlySerCysTyrAlaPheHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db 2834 TTCGGGGTCTCATTTCTATCACCGGCATCATGATATTTTCTATGCTGCTGTTTC 2893
QY 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2894 CGGTTTGGTCTACTTGGTGGCAATAGATTCTGAACCTTTCCAGGATGTTCTTTGGTA 2953
QY 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 2954 TTTCTAGCTATTCTTTGGTGCATGGCAGTGGGCGAGGTCAGTTCAATTGCTCTCTGAC 3013
QY 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlyLysProAsn 962
Db 3014 TATGCCAAGCCAAAGTATCAGCAGCCAGCTCATCATGATCATTTGAAAAGCCCTCTG 3073
QY 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db 3074 ATTCACAGCTACAGCCCTCAGCGCCTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133
QY 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3134 AATGAGGTGCTGTTCACATATCCACATCCAGCACATCCCGCTGCTCCAGGGGCTGAGC 3193
QY 1003 LeuSerIleGluArgGlyLysThrValAlaPheValIleGlySerSerGlyCysGlyLysSer 1022
Db 3194 CTCGAGGTGAGAGGGCCAGACGCTGGCCCTCTGTAGTAGCAGTGGCTGTGGAGAGC 3253
QY 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAsp 1042
Db 3254 ACAGTTGTTTCAGCTCTAGAGCGCTTCTATGACCCCTTGCTGCTGCTTAATTGAT 3313
QY 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db 3314 GGCAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373
QY 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3374 CAGGAGCCCATCTGTTTGACTCGAGATGCGGAGAAATTCCTGATGAGACAACAGC 3433
QY 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3434 CGGGTGTATCATAGAGATTTATCCAGCGAGCCCAAGGAGCCCAATACACCACTTC 3493
QY 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3494 ATCCAGACACTCCCTCAGAAATAACACACAGAGTAGGAGACAAAGAACCCAGCTCTCT 3553
QY 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeu 1142
Db 3554 GGTGGCCAGAAACAGCGCATGTGCATAGCTCGCGCTCTTTGACAGCCCTCATATTTTG 3613
QY 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db 3614 CTTTGGATGAAGCTATCATCAGCTCTGGATACAGAAAGTGAAGGTGTGTCGAAGGCC 3673
QY 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db 3674 CTGGCAAAAGCCAGAGAAGGCGCACCTGTCATTTGTATCGCCCAACCGCTTGTCCACCATC 3733
QY 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db 3734 CAGAATGCAGATTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGGCACAT 3793
QY 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3794 CAACAGCTGCTGCCCAAGAAAGCAHCTATTTTCCATGTGTAGTCTCAGGCT 3847
RESULT 7
AAD03505
```

```

ID AAD03505 standard; cDNA; 4279 BP.
AC AAD03505;
DT 13-JUN-2001 (first entry)
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
OS Canis familiaris.
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT replace (91, T)
FT /*tag= b
FT allele replace (607, C)
FT /*tag= c
XX WO200123540-A2.
PN 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ,
XX WPI; 2001-235373/24.
DR P-PSDB; AAE00309.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 9; Page 93-99; 111pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDRI) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX
SQ Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4279
Score: 3530.50 Matches: 682
Percent Similarity: 73.13% Conservative: 238
Best Local Similarity: 54.21% Mismatches: 269
Query Match: 57.32% Indels: 69
DB: 22 Gaps: 9

US-09-873-409-5 (1-1222) x AAD03505 (1-4279)
```

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu	20
Db	167	ATGTTGGTGGGCAATATGGCTGCCATCATCCATGAGCTGCATCCCTCTCATGATCGT	226
Qy	21	ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeu	34
Db	227	GTTTTGGAAACATGACAGATAGCTTTGCNAATGACAGGAATTTCAAGAAACAAAACCTTTT	286
Qy	35	-----ValGlnThrAenThrTyrSerPhePheArg-----	44
Db	287	CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTTCATCAACCATCTGGAG	346
Qy	45	-----LeuThrLeuTyrTyrValGlyVileGlyValAlaAlaLeuPhe	59
Db	347	GAGGAANTGACACCGTATGCCTATTATTACAGTGGATCGGTGCTGGCGTCTGGTGGCT	406
Qy	60	GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg	79
Db	407	GCTTACATCCAGGTTTCATCTGGTGCCTGGCAGCAGGAAGACAGACATCTCAAAATTTAGA	466
Qy	80	LysGlnPhePheHisSerValIleAlaGlnAspIleGlyTrpPheAspSerCysAspIle	99
Db	467	AAACAATTTTTTTCATGCTATCATGACAGGAGATTTGGCTGGTTGACGTCATGACGCTT	526
Qy	100	GlyGluLeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp	118
Db	527	GGGGAGCTTAACCCCGGCTCAGACAGCATGCTCCAAANTCAATGAAGGAATTTGGCGAC	586
Qy	119	LysIleAlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	587	AAAATTTGGAATGTTCTTTCAATCAATAGCAACATTTTCCCGGTTTTATAGTGGGGTTT	646
Qy	139	ValLysGlyTyrTrpLysLeuThrIleValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	647	ACACGTGGTTTGAAGCTAACCCCTTGATTTTGGCCATCAGCCCTGTTCTTTGGACATTTC	706
Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
Db	707	GCCGCCATCTGGGCAAGAATACATCTCTCAATTACTGATAAAGAACTCTGGCCCTATGCA	766
Qy	179	LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	767	AAAGCTGGACAGTAGCTGAAGAAGTCTTTAGCAGCAATCAGAACTGTGATTTGGCTTTGA	826
Qy	199	AlaGlnGluLysGluIleuGlnArgSerPheLeuLeuAenIleThrArgTyrAlaTrpPhe	218
Db	827	GGACAAAGAAGAAGAACTTGAA-----	847
Qy	219	TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAspLeu	238
Db	848	-----AGTCAACAAAAATTTA	865
Qy	239	LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla	258
Db	866	GAAGAAGCTTAAGGAATTTGGATAAAGAAGCTATCACGGCCACATTTCTATTGGTGCC	925
Qy	259	ValTyrPhePheMetAenGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle	278
Db	926	GCTTTCTATTGATCTATGATCATATGCTCTGGCTTTCTGGTATGGGACCTCTCTGGTC	985
Qy	279	LeuAenGlyGluProGlyTyrThrIleGlyThrValIleAlaValPhePheSerValIle	298
Db	986	CTCTCCAGTGAA-----TATTCTATTGACAAGTACTCCTGCTCTCTTTCTGTATT	1039
Qy	299	HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg	318
Db	1040	ATTGGGGCTTTTAGTATTGGACAGGCATCCCCNAGCATTTGAAGCATTTGCAACGCAAGA	1099
Qy	319	GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer	338
Db	1100	GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTTGACAGCTATTTCG	1159

Qy	339	ThrAlaGlyTyrIysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe	358
Db	1160	AAGAGTGGACATAAACCAGATATATTAAAGGAAAATTTTGAATTCCAAAATGTTCACATTC	1219
Qy	359	AsnTyrProSerArgProSerIleLeuLysGlyLeuAsnLeuArgIleLysSer	378
Db	1220	AGTTACCCTTCTCGAAAAAGATTAGAATCTTTAAAGGGTCTCAACTGAAGGTTCCAGAGT	1279
Qy	379	GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu	398
Db	1280	GGCAGACAGTGGCGTGTTGGGAACAGTGCCTGCCGGAAGAGCACACCGTCGACGCTG	1339
Qy	399	LeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArg	418
Db	1340	ATGCAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGGACAGACATTAGG	1399
Qy	419	AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu	438
Db	1400	ACCATAATGTAAAGGCATCTTCGGGAAATTACTGGTGTGGTAGTCAGAGCCGTGTGTG	1459
Qy	439	PheGlyThrThrIleSerAsnAnenIleLysTyrGlyValArgAspValThrAspGluGlu	458
Db	1460	TITCCCACACGATAGCTGNAACAATTCCTATGGCCGCGAAAATGTCCACATGATGATG	1519
Qy	459	MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys	478
Db	1520	ATTGAGAAAGCTGTTAAGGAAGCCAATGCTCTATGATTTTTATCATGAACCTACCTAATAA	1579
Qy	479	PheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIle	498
Db	1580	TTTGACACTCTGGTTGGAGAGACAGAGGGGCCACGCTGAGTGGTGGACAGAAAACAGAAATC	1639
Qy	499	AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer	518
Db	1640	GCCATTGCTCGGGCCCTGGTTGCCAACCCCAAGATCTCTCTCTCGATGAGCAACGTCA	1699
Qy	519	AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGly	538
Db	1700	GCTCTGACACTGAAAGTGAAGCAGTGGTTTCAGGTGGCCCTGGATAAGCCAGAAAAGGC	1759
Qy	539	ArgThrThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal	558
Db	1760	CGGACTACCAATTGTGATAGCTCATCGTTTGTCTACAGATTGTAATGCCGATGTCAATGCT	1819
Qy	559	ThrLeuLysAspGlyMetLeuAlaGluLysGlyValAlaHisAlaGluLeuMetAlaLysArg	578
Db	1820	GGTTTTGATGATGAGTCAVTTGTGGAGAAGGAATCATGATGAACCTCATGAAGAGGAG	1879
Qy	579	GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu	598
Db	1880	GGCATTCTTCAAACCTGTGCAATGCAG---ACAAGAGGAATGAAATTGAGTTAGAA	1936
Qy	599	SerMetThrTyrSerThrArgLysThrAsnSerLeuProLeuHisSerValLysSer	618
Db	1937	AATGCCACTGCTGAATCCAAAAGTGAAGTGAATGCCTTGGAAATG---TCTCCAAAAGAT	1993
Qy	619	IleLysSerAspPheIle-----	624
Db	1994	TCAGGGTCCAGTTTAAATAAAAAGAAAGATCAACTCGCAGGAGTATACATGCACCAACAGGC	2053
Qy	625	---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu	643
Db	2054	CAAGACAGAAAGCTTGGTACAAAAGAGGACTTCAATGAGAATGTACCTCCAGTTTCCCTTC	2113
Qy	644	LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla	663
Db	2114	TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTATTATTTGTGGTGGTATATTTTGT	2173
Qy	664	SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr	683
Db	2174	GCTATTATAAACGGAGCCCTGCAACACAGCATTTTCAATAATATTTTCAAGGATATATAGG	2233
Qy	684	MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet	702

```

Db 2234 ATCTTTACCGAGATCAGAGTCTCTGAAACAAACGACAGATATAGTAAACATGTTTCTGTGA 2293
QY 703 IlePheValIleLeuGlyValIleCysPheValSerPheMetGlnGlyLeuPheTyr 722
Db 2294 TTGTTTCTAGTCTTGGAAATATTTCTTTATTATACATTTTCTCCAGGGCTTCACATTT 2353
QY 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheIysAlaMetLeu 742
Db 2354 GGCAAGCTGGGGAGATCTCTACATAGACGGCTTCGATACATGTTTTCAGATCCATGCTG 2413
QY 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762
Db 2414 AGACAGATGTCAGCTGGTTTGATGACCTTAAACACCACTGGACATTCACACAGG 2473
QY 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db 2474 CTTGCCAATGATCGCGCTCAAGTTAAAGGGCTATAGTTCCAGGCTGCTGTCTATACC 2533
QY 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db 2534 CAGAATATAGCAAACTCTGGGACAGGCATTTATATATCTTAACTATCTATGCTTGGCAATTA 2593
QY 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrClyMetIleGluThrAla 822
Db 2594 ACATTTTACTCTTGCAATTTGACCAATTTGACCAATAGCAGGAGTTGTTGAATGAAA 2653
QY 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db 2654 ATGTTCTCTGGACACACTGAAGATAGAAAGACTAGAGGAGCTGGGAAGATTGCT 2713
QY 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db 2714 ACAGAAGCCATCAAAACTTCGAACTGTGTTCTTTGACTCGGAGCAGAAAGTTTGA 2773
QY 863 GlnMetTyrGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db 2774 TCATGTATGACACAGAGTTGCAAGTATCCATACAGAAACTCTTTGAGGAAAGCACACATC 2833
QY 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db 2834 TTCGGGGTCTCATTTCTATCACCCAGGCATATGATATTTTCTATGCTGCTGTTTC 2893
QY 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db 2894 CGGTTGGTGCTACTTGGTGGCAAAATGAGTTCTATGAACCTTTCAGGATGTTCTTTGGTA 2953
QY 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db 2954 TTCTCAGCTATTGCTTTGTTGTCATGGCAGTGGGCGAGTTCAGTTTCATTGCTCCTGAC 3013
QY 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsn 962
Db 3014 TATGCCAAGCCAAAGTATCAGACGCCAGCTCATCATGATCAITGAAAGCCCTCTG 3073
QY 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db 3074 ATTGACAGCTACAGCCCTCACGGCTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133
QY 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db 3134 AATGAGGTGCTGTTCAACTATCCACTCGACACAGACATCCCCGTGCTCCAGGGGCTGAGC 3193
QY 1003 LeuSerIleGluArgGlyLysThrValAlaPheValClySerSerGlyCysGlyLysSer 1022
Db 3194 CTGAGGTGAAGAAAGGCCAGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3253
QY 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db 3254 ACAGTTGTTTCAGCTCTAGAGCGCTTCTATGACCCCTTGCTGCTGCTGCTGCTGCTGCTGCT 3313
QY 1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062

```

```

Db 3314 GGCAAGAGATAAAGCACCCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373
QY 1063 GlnGluProValIlePheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db 3374 CAGGAGCCCATCTCTGTTGACTGCAGCATTCGCGAACAATTCCTTATGGAGACAACAGC 3433
QY 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db 3434 CGGGTCGTATCATCATGAAGAGATTATGCGGCGAGCCCAAGAGGCCAATACACCACTTC 3493
QY 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3494 ATCGAGACACTCCCTGAGAAATACACACACAGTAGGAGACAAAGAACCCAGCTCTCT 3553
QY 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db 3554 CGTGGCCAGAAACAGCGCATTCGCATAGCTCGCGCTCTTTGTAGACAGCTCATATTTTG 3613
QY 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db 3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAGCC 3673
QY 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db 3674 CTGGCAAAAGCCAGAGAGCGCCGACCTGCTATGTGATCCGCCACCGCTTGTCCACCATC 3733
QY 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db 3734 CAGAATGCAGATTTAATAGTGTGTTTCAGATGGCAAGTCAAGGAGCATGGCACACAT 3793
QY 1203 GlnGluLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3794 CAACAGCTGCTGCCAGAAAGGCATCTATTTTCCATGTCAGTGTCCAGGCT 3847

RESULT 8
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
Key Location/Qualifiers
CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX
PN WO200123540-A2.
XX
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI; 2001-235373/24.
DR P-PSDB; AAE00303.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX

```



Claim 3; Page 58-63; 11lpp; English.

PS The invention relates to dog P-glycoprotein (PGP) also referred  
 XX as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterized by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) cDNA. This  
 CC sequence is also referred as genotype C cDNA. The  
 CC PGP enzyme functions as an efflux pump exporting small molecules  
 CC across the cell membrane. This enzyme is a member of the ABC  
 CC transporter family.

XX Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

# Alignment Scores:

Pred. No.:	0	Length:	4279
Score:	3525.50	Matches:	681
Percent Similarity:	73.05%	Conservative:	238
Best Local Similarity:	54.13%	Mismatches:	270
Query Match:	57.24%	Indels:	69
DB:	22	Gaps:	9

US-09-873-409-5 (1-1222) x AAD03488 (1-4279)

Qy	1	MetLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	167	ATGTTGGTGGGACAATGGCTGCATCATCCATGGAGCTGCACCTCTCATGATCGT	226
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu	34
Db	227	GTTTTGGAAACATGACAGATAGCTTTGCAATGCAGGAATTTCAAGAACAAACTTTT	286
Qy	35	-----ValGlnThrAsnThrTyrSerPhePheArg-----	44
Db	287	CCAGTTATAATTAAGAGTATTACGAACAATACACAACTTTTCATCAACCATCTGGAG	346
Qy	45	-----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe	59
Db	347	GAGGAATGACCCGATATGCTATTATTACAGTGGGATCGGTGCTGGTGGCT	406
Qy	60	GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg	79
Db	407	GCTTACATCCAGGTTTCATTCGTGGCTGGCAGGAGGACAGACTACTCAAAATTAGA	466
Qy	80	LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle	99
Db	467	AAACAAATTTTTCATGCTATATCGACAGGAGATGGCTGGTTGACGTGCATGAGTT	526
Qy	100	GlyGluLeuAsnThrArgMetThr-----AspIleAspIlyIleSerAspGlyIleGlyAsp	118
Db	527	GGGAGAGCTTAACACCGGCTCACAGACGATGCTCCAAATCAATGAAGGAATTTGGCGAC	586
Qy	119	LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu	138
Db	587	AAATTTGGAATGTTCTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT	646
Qy	139	ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer	158
Db	647	ACAGTGGTTGGAAGCTAACCTTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTCA	706
Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
Db	707	GCCGCCCATCTGGGCAAGATACTATCTTCTTACTATGAAGAAGTCTTGGCCCTATGCA	766

Qy	179	LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	767	AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAAGCTGTGATTGGCTTTGGA	826
Qy	199	AlaGlnGlyLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe	218
Db	827	GGCAAAAGAAAGAACTTGA-----	847
Qy	219	TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu	238
Db	848	-----AGGTACAACAAAATTTA	865
Qy	239	LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla	258
Db	866	GAAGAAGCTAAAGGAATTTGGATATAAGAAAGCTATCACGCCCAACATTTCTATTGGTGCC	925
Qy	259	ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIle	278
Db	926	GCCTTCTTATTGATCATGATCATATGCTCTGGCTTTCTGGTATGGACCTCTCTGGTC	985
Qy	279	LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle	298
Db	986	CTCTCAGTGAA-----TATTCTATTGGACAAGTACTCACTGCTCTCTTTCTGTATTA	1039
Qy	299	HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg	318
Db	1040	ATTGGGGCTTTTATGATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA	1099
Qy	319	GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer	338
Db	1100	GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTTGACACTATTTCG	1159
Qy	339	ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe	358
Db	1160	AAGAGTGGACATTAACAGCATATATATTAAGGAAATTTGGAATTCAAAATGTTCACTTC	1219
Qy	359	AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer	378
Db	1220	AGTTACCTTCTCGAAAAGAAAGTAAAGATCTTAAAGGCTCTCAACCTGAAAGTTACAGT	1279
Qy	379	GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu	398
Db	1280	GGGCAGACAGTGGCTGGTGGGAAACAGTGGCTGGGGAAGAGCAGCCGTCGACGTG	1339
Qy	399	LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg	418
Db	1340	ATGCAGAGGCTCTATGACCCACAGATGGATGGTCTGTATTGATGACAGGACATTAGG	1399
Qy	419	AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu	438
Db	1400	ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGTGTGGTGGTGCAGGAGCCTGTGTTG	1459
Qy	439	PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu	458
Db	1460	TTTGCCACACAGATAGCTGAAACATTTGCTATGGCCGCGGAAATGTCACCATGATGAG	1519
Qy	459	MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys	478
Db	1520	ATTGAGAAAGCTGTTAAGGAAGCCATGCTCTATGATTTTATCATGAACCTACCTAATAA	1579
Qy	479	PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle	498
Db	1580	TTTGACACTCTGGTTGGAGAGAGAGGGGGCCAGCTGAGTGGTGACAGAGAACAGAGANTC	1639
Qy	499	AlaIleAlaAlaLeuValArgAsnProIleIleLeuLeuAspGluAlaThrSer	518
Db	1640	GCCATTGCTCGGGCCCTGGTTTCGCAACCCCAAGATCTCTCTGCTGGATAGGCAACGTC	1699
Qy	519	AlaLeuAspSerGluSerLysSerValGlnAlaAlaLeuGluLysAlaSerLysGly	538
Db	1700	GCTCTGACACTGAAAGTGAAGCAGTGGTTGAGTGGCCCTGGATAGGCCAGAAAGGC	1759
Qy	539	ArgThrThrIleValValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuVal	558



|||||  
Db 1760 CGGACTACCATTTGATAGTCATCGTTGCTACAGTTTCGTAAATCCGATGTCATTGCT 1819  
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyValAlaHisAlaGluLeuMetAlaLysArg 578  
|||||  
Db 1820 GGTTTTGGATGATGAGTCACTTGTGGAGAAAGAAATCATCATGCAATCATGAAAGAGAG 1879  
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598  
|||||  
Db 1880 GGCATTTTACTTCAAACTTGTCAAAATGCAG--ACAAGAGAGAAATGAAATGAGTTAGAA 1936  
Qy 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
|||||  
Db 1937 AATGCCACTTGGTAATCCAAAGTGAAGTGAATGCTTGGAAATG---TCTCAAAGAT 1993  
Qy 619 IleLysSerAspPheille----- 624  
Db 1994 TCAGGGTCCAGTTTAAATAAAAGAGATCAACTCGCAGGAGTATATACATGCCACCAAGGC 2053  
Qy 625 ----AspLysAlaGluSerThrGlnSerLysGluLysLeuSerLeuProGluValSerLeu 643  
|||||  
Db 2054 CAAGACAGAAAGCTTGGTACAAAGAGAGACTTGAATCAGAAATGTACCTCCAGTTCCCTTC 2113  
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 663  
|||||  
Db 2114 TGGAGGATTTCTGAAGCTGAACCTCAACTGAATGGCCCTTATTTTGTGGTATATTTTGT 2173  
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683  
|||||  
Db 2174 GCTATTTAATAACGGAGGCTGCACACAGCATTTTCAATATATATTTTCAAGGATATAGGG 2233  
Qy 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluLysIleTyrSerMet 702  
|||||  
Db 2234 ATCTTTACCCGATCAGGATCCTGAAACAAAACGACAGAAATAGTAACATGTTTCTGTA 2293  
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722  
|||||  
Db 2294 TTGTTTCTAGTCTCGGAATATTTCTTTTATTACATTTTCTCCAGGGCTTTCACATTT 2353  
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742  
|||||  
Db 2354 GCGAAGCTGGGAGATCTCACTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTG 2413  
Qy 743 TyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762  
|||||  
Db 2414 AGACAGGATGTCACTGCTGTTGATGCCCTTAAACACCACTGGAGCATTTGACCAACAGG 2473  
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782  
|||||  
Db 2474 CTGCAATGATCGGCTCAAGTTAAAGGGGCTATAGTTCCAGGCTTGTGTCATTACC 2533  
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrTpGluMet 802  
|||||  
Db 2534 CAGATATAGCAAACTTGGGACAGCATTTATATCTTAATCTATGTTGGCAATTA 2593  
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822  
|||||  
Db 2594 ACATTTTACTCTTACCAATGTGACCATCATTTGCAATAGCAGAGTGTGTGAAATGAAA 2653  
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842  
|||||  
Db 2654 ATGTTCTCTGGCAAGCACTGAAAGATAGAAAGAGCTAGAGAGGCTGGAAGATGCT 2713  
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862  
|||||  
Db 2714 ACAGAGCCATCGAAACCTTCCGAATGTTGTTCTTTCAGCTCGGAGCAGAGTTTGA 2773  
Qy 863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882  
|||||  
Db 2774 TACATGTATGCACAGAGTTTGCAGATACCATACAGAAATCTTTTGGAGAGACACATC 2833  
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902  
|||||

2834 TTCCGGGTCTCATTTTCTATACCCAGGCAATGATGATATTTTCTATGCTGCTGCTTTTC 2893  
Qy 903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922  
|||||  
Db 2894 CGTGTGGTGCCTACTTGGTGCAAAATGATTTCACTGAACCTTTTCAGGATGTTCTTTTGGTA 2953  
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942  
|||||  
Db 2954 TTCTCAGCTATTGTTCTTGTGTCATGGCAGTGGGCGAGCTCAGTTTCATTGCTCTCTGAC 3013  
Qy 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsn 962  
|||||  
Db 3014 TATGCCAAAGCCAAAGTATCATGAGCCCACTGATCATGATCATTTGAAAAGCCCTCTG 3073  
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982  
|||||  
Db 3074 ATTGACAGCTACAGCCCTCAGCGCCTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133  
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002  
|||||  
Db 3134 AATGAGGTGCTGTTCAACTATCCCACTCGACACGACATCCCGTGTCTCCAGGGCTGAGC 3193  
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022  
|||||  
Db 3194 CTCGAGGTGAAGAGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGAAAGAGC 3253  
Qy 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042  
|||||  
Db 3254 ACAGTTGTTTCAGTCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTCTAATTGAT 3313  
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValPro 1062  
|||||  
Db 3314 GCGAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373  
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082  
|||||  
Db 3374 CAGGAGCCCATCTGTTGACTGCGACATTTGCCGAAACATTTGCCCTATGGAGAACACAGC 3433  
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102  
|||||  
Db 3434 CGGTCGTATCATCATGAGAGATTATGAGGCGAGCCCAAGGAGGCCAATACACCACTTC 3493  
Qy 1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122  
|||||  
Db 3494 ATCGAGACACTCCCTGAGAAATACAAACACAGATAGGAGACAAAGGAACCCAGCTCTCT 3553  
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeu 1142  
|||||  
Db 3554 GGTGGCCAGAAACAGCGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3613  
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162  
|||||  
Db 3614 CTTTGGATGAGCTACATCAGCTCTGATACAGAAAGTGAAGAGTTGTTCCAGAGGCC 3673  
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182  
|||||  
Db 3674 CTGGCAAAAGCCAGAGAGCGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3733  
Qy 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202  
|||||  
Db 3734 CAGAAATGCGAGATTTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGGCACAT 3793  
Qy 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
|||||  
Db 3794 CAACAGCTCTGCCCGAAGAGGCATCTATTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3847  
RESULT 9  
AAD03489  
ID AAD03489 standard; cDNA; 4317 BP.  
XX  
AC AAD03489;  
XX  
DT 13-JUN-2001 (first entry)

```

XX DE Dog P-glycoprotein (PGP) cDNA #2.
XX DE
XX KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX KW drug bioavailability; transgenic animal; genetic model; ss.
XX OS
XX OS Canis familiaris.
XX FH Location/Qualifiers
XX FT 70..3912
XX FT /*tag= a
XX FT /product= "Dog P-glycoprotein (PGP) #2"
XX PN WO200123540-A2.
XX XX
XX PD 05-APR-2001.
XX XX
XX PF 28-SEP-2000; 2000WO-US26767.
XX PR 28-SEP-1999; 99US-0156510.
XX XX
XX PA (GENT-) GENTEST CORP.
XX
XX PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX XX
XX DR WPI: 2001-235373/24.
XX DR P-PSDB; AAE00304.
XX XX
XX PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX PT for determining the bioavailability of drugs and for screening for dog
XX PT PGP inhibitors -
XX XX
XX PS Claim 1; Page 66-72; 111pp; English.
XX
XX CC The invention relates to dog P-glycoprotein (PGP) also referred
XX CC as multidrug transporter (MDR1) and nucleic acids encoding them.
XX CC The invention also includes fragments and biologically functional
XX CC variants of dog P-glycoprotein. PGP and their nucleic acids are
XX CC useful for determining the bioavailability of drugs and for
XX CC screening PGP inhibitors. They are useful for the diagnosis and
XX CC treatment of conditions characterised by PGP activity, by
XX CC reducing or increasing PGP activity in a cell. PGP nucleic acids
XX CC are used as oligonucleotide probes. Complements of PGP nucleic
XX CC acids are useful as antisense oligonucleotides, to induce a PGP
XX CC 'knockout' phenotype. They are used to prepare a non-human
XX CC transgenic animal, which are valuable as genetic models for
XX CC human diseases.
XX CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
XX CC PGP enzyme functions as an efflux pump exporting small molecules
XX CC across the cell membrane. This enzyme is a member of the ABC
XX CC transporter family.
XX
XX SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4317
Score: 3525 50 Matches: 680
Percent Similarity: 73.05% Conservative: 239
Best Local Similarity: 54.05% Mismatches: 270
Query Match: 57.24% Indels: 69
DB: 22 Gaps: 9

US-09-873-409-5 (1-1222) x AAD03489 (1-4317)
QY 1 MetLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 217 ATGTGGTGGGACAAATGGCTGCATCATCGAGCTGCATCCCTCTCATGATGCTG 276
QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeu----- 34
Db 277 TTTTGGAAACATGACAGATAGCTTTGCAAAATGACGAATTTCAAGAAACAAACTTTT 336
QY 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44

```

```

Db 337 CCAGTTTATAATTAAGAGTATTACGAAACAATACACAACATTTCAACCATCTGGAG 396
QY 45 -----LeuThrLeuTyrValGlyValGlyValAlaAlaLeuPhe 59
Db 397 GAGGAAATGACCAGTATGCCCTATTATTACAGTGGATCGGTGCTGGTGGTGGT 456
QY 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 457 GCTTACATCCAGTTTCATTCTGGTGGCAGCAGGAGGAGACAGATCTCAAAATTAGA 516
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 517 AAACAATTTTTCATGCTATCATCGACAGAGATGGCTGTTTGACGTGATGACGTT 576
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 577 GGGGAGCTTAACACCCGGCTCACAGCATGTCTCCAAATCAATCAAGGAAATGGCGAC 636
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 637 AAAGTTGGATGTTCTTTCAATCAATAGCAACATTTTACCAGTTTATAGTGGGTTT 696
QY 139 ValLysGlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 697 ACACCTGGTTGGAAGCTAACCCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACATTC 756
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 757 GCGCCATCTGGGCAAGATATCTATCTTATTACTGATAAAGAACTCTGGCCTATGCA 816
QY 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 817 AAAGCTGGAGCAGTAGCTGAGAGAGCTTAGCAGCAATCAGAACTGATGCTTGGCTTGA 876
QY 199 AlaGlnGlyLeuGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe 218
Db 877 GGACAAAGAAAGAACTTGAA----- 897
QY 219 TyrPheProGlnThrLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db 898 -----AGGTACAAACAAAATTTA 915
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 916 GAAGAAGCTAAAGAATTGGGATAAAGAAGCTATCACGCCCAACATTTCTATGTGTGCC 975
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278
Db 976 GCTTCTTATTGATCATCATCATATGCTCTGGCTTTCTGGTATGGGACCTCTTGCTC 1035
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIle 298
Db 1036 CTCTCCAGTGA-----TATACTATTGGACAGGTACTCACTGCTCTTCTTTCTGATTA 1089
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1090 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1149
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1150 GGAGCAGCTTATGAAATCTTCAAGATAATTCACATAATACCAAGCATTTGACACTATTTCG 1209
QY 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
Db 1210 AAGAGTGGACATAAACACAGATAATATTAAGGGAATTTGGAATTTCAAAATGTTACTTC 1269
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1270 AGTTACCTTCTCGAAAGAGAGTATAGATCTTAAGGGTCTCAACCTGAAGGTTACAGT 1329
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398

```

Db 1330 GGGCAGACAGTGGCGTGGTGGGAACAGTGGCTGGGGAAGACGACGACCGTGCAGCTG 1389  
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
Db 1390 ATGCAGAGGCTCTATGACCCACAGATGCATGGTCTGTATTGATGCAGCAGCATAGG 1449  
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
Db 1450 ACCATAAATGTAAGGCATCTTCGGGAATATTACTGTGTGTGGTGCAGTCAGGAGCGCTGTGTTG 1509  
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458  
Db 1510 TTTCGCCACCATGATGCTGAACAATTCCTATGCGCGGAAATGTCACCATGGATGAG 1569  
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
Db 1570 ATTGAGAAAGCTGTTAAGGAAGCCATGCTATGATTTTATCATGAACACTACCTAATAA 1629  
Qy 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
Db 1630 TTTGACACTCTGCTGGAGAGAGGGCGCGCTGAGTGGTGGACAGAAACAGAGAATC 1689  
Qy 499 AlaIleAlaArgAlaValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
Db 1690 GCCATTGCTCGGCGCTGCTGTCGAACCCAGATTTCTTCTGCTGGATGAGGCAACGTCA 1749  
Qy 519 AlaLeuAspSerGlnSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538  
Db 1750 GCTCTGGACATGAAGTGAAGCAGTGGTTCAGTGGCCCTGGATGAAGCCAGAAAGGC 1809  
Qy 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspIleVal 558  
Db 1810 CGGACTACCATTTGATAGTCTATCGCTTGTCTACAGTTCTGTAATCCGATGCTCATTTGCT 1869  
Qy 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578  
Db 1870 GGTTTTGATGGATGATCAATGTGGAGAAGGAAATCATGATGAATCAAGAAAGAG 1929  
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598  
Db 1930 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAATGAATTTGATTAGAA 1986  
Qy 599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
Db 1987 AATGCCACTGGTGAATCCAAAGTGAAGATGATGCTTGGAAATG---TCTCCAAAGAT 2043  
Qy 619 IleLysSerAspPheIle----- 624  
Db 2044 TCAGGCTCCAGTTTATAAAAGAAAGATCAACTCGCAGGAGTATACATGCACCAAGGC 2103  
Qy 625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643  
Db 2104 CAAGCAGAAAGCTTGGTACAAAGAGGACTTGAATGAGATGTTCTTCAGTTTCTTTC 2163  
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAla 663  
Db 2164 TGGAGATTCTGAAGCTGAACCTCACTGAATGCGCTTATTTCTGTGTGATATTTTGT 2223  
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683  
Db 2224 GCTATTATAACGGCGCTGCAACCCAGCATTTTCAATAATATTTTCAAGGATTAAGG 2283  
Qy 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702  
Db 2284 ATCTTTACCGAGATGAGGATCTCTGAACCAAAACGACAGAAATAGTAACATGTTTCTGTA 2343  
Qy 703 IlePheValIleLeuGlyValIleLysPheValSerThrPheMetGlnGlyLeuPheTyr 722  
Db 2344 TTGTTTCTAGTCTCTGGAATTAATTTCTTTTATATACATTTTCTCCAGGCGCTTCACATTT 2403  
Qy 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742  
Db 2404 GGCAAGCTGGGAGATCTCTCACTAAGCGGCTTCGATACATGTTTTCAGATCCATCTTC 2463

Qy 743 TyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762  
Db 2464 AGACAGGATGTGAGTGGTGTTCATGACCCCTAAATAAACACCACTGGAGCATTTGACACAGG 2523  
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuLeu 782  
Db 2524 CTTGGCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCAGGCTTCTGTCATTACC 2583  
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMet 802  
Db 2584 CAGATATAGCAAACTCTGGACAGGCATTTATATCTTAATCTATGTTGGCAATTAT 2643  
Qy 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822  
Db 2644 ACACATTTTACTCTTAGCAATTTGATCCCATCATTCGAATAGCAGGAGTTCTTGAATGAAA 2703  
Qy 823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842  
Db 2704 ATGTTGTCTGGCAAGCACTGAAGAAGATAGAAAGAGCTAGAAAGAGCTGGGAAGATTGCT 2763  
Qy 843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862  
Db 2764 ACAGAAGCCATCGAAACTTCCGAATCTGTTGTTCTTTGACTCGGAGGCAGAAGTTTGA 2823  
Qy 863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882  
Db 2824 TACATGTATGCACAGATTTGCAAGTACCATACAGAAACTCTTTGAGGAAGACACATC 2883  
Qy 883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaGlyPhe 902  
Db 2884 TTCGGGCTCTCAATTTCTATCACCCAGCAATGATGATTTTTCCTATGCTGGCTGTTTC 2943  
Qy 903 ArgPheGlyValaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922  
Db 2944 CGTTTGTGGCTCTCTTGTGCAATGATGTTCACTGCAACTTTTCAGGATGTTCTTTTGGTA 3003  
Qy 923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942  
Db 3004 TTCTCAGCTATTGCTTCTTGTGCTGATGCGAGTGGGCGAGTTCAGTTTCCTCTGAC 3063  
Qy 943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsn 962  
Db 3064 TATGCCAAAGCCAAAGTATCAGCAGCCCACTCATCATGATCATTTGAAAAAGCCCTCTG 3123  
Qy 963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982  
Db 3124 ATTGACAGCTACAGCCCTCAGGCTCAAGCCCAATACGTTTGAAGAAATGTGACATTT 3183  
Qy 983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002  
Db 3184 AATGAGTCTGTTCAACTATCCACTCCAGCAGACATCCCGTGTCTCCAGGCGGTGAGC 3243  
Qy 1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022  
Db 3244 CTCAGGTGAAGAAGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3303  
Qy 1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042  
Db 3304 ACAGTTGTTGAGTCTCAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTCTAATTGAT 3363  
Qy 1043 GlyValAspAlaLysGluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValPro 1062  
Db 3364 GGCAAGAGATAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3423  
Qy 1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082  
Db 3424 CAGGAGCCCATCTGTTTCACTGCAGCATTTGCCGAAACATTTGCCCTATGGAGACAACAGC 3483  
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102  
Db 3484 CGGTCGTATCATGTAAGAGATTATGTCAGGCGGCCAAGAGGCCCAACATACACCACTTC 3543

QY 1103 ileGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122  
 DB 3544 ATCGAGACACTCCCTGAGAAATACACACAGAGTAGGAGACAAAGAACCCAGCTCTCT 3603  
 QY 1123 GlyGlyGlnLysGlnArgLeuAlaAlaAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142  
 DB 3604 GGTGGCCAGAAACAGCGCATTTGCGCATAGCTCGCGCTCTTGTAGACGCGCTCATATTTG 3663  
 QY 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGluLysValValGlnHisAla 1162  
 DB 3664 CTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAGGTTGCCAAGGCC 3723  
 QY 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182  
 DB 3724 CTGGACAACCCAGAGAGGCCGCGACCTGATGTGATCGCCACCGCTTGTCCACCATC 3783  
 QY 1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202  
 DB 3784 CAGAATGCAGATTTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGCGCACAT 3843  
 QY 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPhelLysLeuValAsnAlaGlnSer 1220  
 DB 3844 CAACAGCTGCTGCTCAGAAAGGCACTATTTTCCATGATCAGTGTCACGCT 3897  
 RESULT 10  
 AAF86127  
 ID AAF86127 standard; cDNA; 4186 BP.  
 XX AAF86127;  
 AC AAF86127;  
 XX AAF86127;  
 DT 25-JUN-2001 (first entry)  
 XX Cynomologous monkey P-glycoprotein cDNA.  
 DE Cynomologous monkey P-glycoprotein cDNA.  
 XX Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;  
 KW efflux pump; ss.  
 XX Macaca fascicularis.  
 OS  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 100..3942  
 FT /\*tag= a  
 FT /product= "PGP"  
 FT /note= "P-glycoprotein"  
 XX  
 PN WO200123565-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 XX 28-SEP-2000; 2000WO-US26592.  
 XX  
 XX 28-SEP-1999; 99US-0156921.  
 PR 12-OCT-1999; 99US-0158818.  
 XX  
 XX (GENT-) GENTEST CORP.  
 PA  
 XX  
 XX Stocker PJ, Steimel-Crespi DT, Crespi CL;  
 PI  
 XX  
 DR WPI; 2001-316136/33.  
 DR P-PSDB; AAB81064.  
 DR  
 XX Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 PT cell -  
 XX  
 XX Example 1; Page 51-57; 84pp; English.  
 PS  
 XX  
 XX This invention relates to a polynucleotide sequence encoding a  
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention

CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents cDNA encoding  
 CC cynomologous monkey P-glycoprotein.  
 XX  
 SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 4186  
 Score: 3524.50 Matches: 679  
 Percent Similarity: 73.75% Conservative: 248  
 Best Local Similarity: 54.02% Mismatches: 261  
 Query Match: 57.23% Indels: 69  
 DB: 22 Gaps: 11  
 US-09-873-409-5 (1-1222) x AAF86127 (1-4186)  
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20  
 DB 250 ATGCTGCTGGAACTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33  
 DB 310 GTGTTTGGAGACATGACGATACCTTTTGCAATGTCAGGAAATTTAGGAGATTTAGGAGCT 369  
 QY 34 LeuValGlnThrAsnThrTyrSer-----41  
 DB 370 CTGTTG---ACTAATAGCAGTAAATATCATGTACATGCCCCCTCATGATCTGGAGGAA 426  
 QY 42 ---PhePheArgLeuThrLeuTyrValGlyIleGlyValaAlaAlaLeuIlePheGly 60  
 DB 427 GATATCACCAGGTATCGCTATTATTACAGTGGAAATGGTCTGGGGTGTGGTGTCTGCT 486  
 QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
 DB 487 TACATTTCAGTTTCATTTCGTGCTGGCAGCTGGGAAGACAAATACACAAATATAGAAA 546  
 QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
 DB 547 CAGTTCATGCTATAATGCGCAGGAGTAGGCTGTTGATGTGACGATGTTGGG 606  
 QY 101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119  
 DB 607 GAGCTTAACACCGCGCTTACAGATGATGCTCCAGATTAATGAGGAATTTGTGACAAA 666  
 QY 120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139  
 DB 667 ATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGTATTATAGTAGGATTTACA 726  
 QY 140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159  
 DB 727 CGTGGTTTGAAGCTAAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTGGACTGTACGT 786  
 QY 160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179  
 DB 787 GCAGCTGGGCAAGATACCTGCTTCATTACTGATAAAGAACTCTTAGCTTATGCAAAA 846  
 QY 180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199  
 DB 847 GCTGGAGCAGTAGCTGAAGAGGCTTGGCAGCAATAGAACTGTGATTCATTTGGAGGA 906  
 QY 200 GlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyr 219  
 DB 907 CAAAAGAAAGAACTCGAA-----924  
 QY 220 PheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLys 239  
 DB 925 -----AGGTACACACAAATTTAGAA 945  
 QY 240 AspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaVal 259

Db 946 GAAGCTAAAGAAATTGGGATAAAGAAAGCTATTACAGCCAAATATTTCTATAGGTGCTGCT 1005  
Qy 260 TyrPhePheMetAsnGlyTyrGlyLeuAlaPheTyrGlyThrSerLeuLeuLeu 279  
Db 1006 TTCTCCTTATCTATGCACTTATGCTCTGGCTTCTGATGGAGCACCTTGGTCTC 1065  
Qy 280 AsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHis 299  
Db 1066 TCAAAAGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAAT 1119  
Qy 300 SerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly 319  
Db 1120 GGGGCTTTTAGTGTGGACAGGCATCTCCAAAGCATTTGAAGCATTTGCAATGCAAGAGGA 1179  
Qy 320 AlaAlaPheHisIlePheGlnValIleAspLysPheProSerIleAspAsnPheSerThr 339  
Db 1180 GCAGCTTTTGAATCTTCAAGATAATTGATAATAAGCCAAATTTGACAGCTATTGCAAG 1239  
Qy 340 AlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsn 359  
Db 1240 AGTGGGCACAAACCATTAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGT 1299  
Qy 360 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 379  
Db 1300 TACCCATCTCGAAAGAAAGTAAAGATCTTGAAGGGCTCGAACCTGAAGGTGCAGAGTGGG 1359  
Qy 380 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 399  
Db 1360 CAGACGGTGGCCCTGGTGTGGAACACAGCGCTGTGGAGAGACAAACGGTCCAGCTGATG 1419  
Qy 400 GlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAla 419  
Db 1420 CAGAGGCTTTATGACCCACAGAGGCGCATGGTCAGTGTGATGCAGCATATTAGGACC 1479  
Qy 420 LeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPhe 439  
Db 1480 ATAAACGTAAAGTGTCTACGGGAAATCATCGGTGTGGTGAAGATGTCACCATGGATGAT 1539  
Qy 440 GlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMet 459  
Db 1540 GCCACCATGATGCTGAAGACATTCCTATGTGTGCTGGAAGATGTCACCATGGATGAT 1599  
Qy 460 GluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPhe 479  
Db 1600 GAGAAAGCTCTCAAGAGACCAATGCTATGACTTTATCATGAACCTGCTCAGAAATTT 1659  
Qy 480 AsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAla 499  
Db 1660 GACACCTGCTGGAGAGAGGGGCCAGCTGAGTGTGGGAGAGAGCAGAGGATCGCC 1719  
Qy 500 IleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAla 519  
Db 1720 ATTGCACGTGCCCTGTTGCAACCCCAAGATCCTCTGCTGACGAGGCGACGTGAGCC 1779  
Qy 520 LeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArg 539  
Db 1780 TTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGCTCTGGATAAGGCCAGAAAGGTCCG 1839  
Qy 540 ThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThr 559  
Db 1840 ACCACCATTTGATAGCTCATCGTTTGTCTACGGTTGCTGAATCCGACGTCATCGTGGT 1899  
Qy 560 LeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGly 579  
Db 1900 TTCGATGATGGAGCTATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAGAGAAAGGC 1959  
Qy 580 LeuTyrTyrSerLeuValMetSerGlnAspIle-----LysLys 592  
Db 1960 ATTTACTTCAAACTTGTCAAAATTCGACAGCAGGAAATGAAATTTAGAAAATGCA 2019  
Qy 593 AlaAspGlu-----GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsn 609  
Db 593 AlaAspGlu-----GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsn 609

Db 2020 GCTGATGAATCCAAAAGTGAAATTTGATACCTTGGAAATGCTCTTCACATGATTCAGATCC 2079  
Qy 610 SerLeu-----ProLeuHisSerValLysSerIleLysSerAspPheIle 624  
Db 2080 AGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTCCGTGGATCATCAAGGCCAA----- 2133  
Qy 625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644  
Db 2134 GACAGAAAGCTTAGTACCAAGAGGCTCTGGATGAAGATATACCTCCAGATTTCTTTTGG 2193  
Qy 645 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 664  
Db 2194 AGGATTATGAAGCTAAATTTAACTGAGTGGCTTATTTTGTGTGTTGTTGTTGTTGTTG 2253  
Qy 665 ValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThrMet 684  
Db 2254 ATTATAAATGGAGGTCTGCAACACGAGCATTTGCAGTAATATTTTCAAGATTTATAGGAT 2313  
Qy 685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703  
Db 2314 TTTACAAGAAATGATGATGCCGAAACCAACACAGCAATAGTAACCTTGTTTTCACTATTG 2373  
Qy 704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723  
Db 2374 TTTCTAGTCTTGGATTTGTTTCTTTTATTACATTTTCTTTCCTTCAGGCTTCACATTTGSC 2433  
Qy 724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743  
Db 2434 AAAGCTGGAGAGATCTCACCAGCGCTCCGATACATGTTTTCGATCCATGCTCAGA 2493  
Qy 744 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763  
Db 2494 CAGGATGTGAGCTGGTTGATGCCCTAAAAACACCACTGGAGCATTTGACTACAGGCTC 2553  
Qy 764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783  
Db 2554 GCCATGATGCTGCTCAAGTTAAGGGCTATAGTTCCAGGCTGCTATATATACCCAG 2613  
Qy 784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 803  
Db 2614 AATATAGCAAACTTTGGGACAGGAATAATTATCTTAATCTATGTTGGCAACTGACA 2673  
Qy 804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823  
Db 2674 CTGTACTCTTAGCAATTTGATCCCATCATTTGCAATAGCAGGAGTTGTGAAATGAAAATG 2733  
Qy 824 MetThrGlyPheAlaAsnLysAspLysGlnGluLysHisAlaGlyLysIleAlaThr 843  
Db 2734 TTGCTGGACAAGCACTGAAGATGAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACT 2793  
Qy 844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863  
Db 2794 GAAGCAATAGAAAACCTTCGAACCTGTTGTTTCTTTGACTCAGGAGCAGAAAGTTTGAACAT 2853  
Qy 864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883  
Db 2854 ATGTATGATCAGAGTTTGCAGGTACCATACAGAAACTCTTTTGAGGAAAGCACACATCTTT 2913  
Qy 884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903  
Db 2914 GGAATCACCTTTTCTTCACGCGCAATGATGATTTTTTCTATGCTGATGATTTTCCCG 2973  
Qy 904 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923  
Db 2974 TTTGGAGCTTACTTTGGTGACACATAGTCTCATGAGCTTTTGAGGATGTTCTGTAGTATTT 3033  
Qy 924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943  
Db 3034 TCAGCTGTTGTTTGGTCCCATGGCGGTGGGCAAGTCAGTTTCATTGCTCTGACTAT 3093  
Qy 944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 963  
Db 3094 GCCAAAGCCAAAGTATCAGCAGCCCATCATCATCATCATTTGAAATTTGATT 3153

```
QY 964 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983
DB 3154 GACAGCTACAGCAGAGAGCCCTAAGCCGAAACACATTGGAGGAATGTCATTTAAT 3213
QY 984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003
DB 3214 GAAGTTGATTCAACTATCCACCGACTGCACATCCAGTGTTCAGGGCTGAGCGCTG 3273
QY 1004 SerIleGluArgGlyThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023
DB 3274 GAAGTGAAGAGGCCAGACGCTGGCCCTGTGGCGGAGAGTGGCTGGGAAGACGACG 3333
QY 1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGly 1043
DB 3334 GTGGTCAGCTCTCTGGAGCGGTTCTATGACCCCTTGGCGGGAGAGTGTCTGTGACGGC 3393
QY 1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063
DB 3394 AAGAAATAAAGCAACTGAATGTTCAAGTGCTCCGAGCACACCTGGGCATCGTGTCCAG 3453
QY 1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083
DB 3454 GAGCCCATCTCTGTTGACTCGACCATTTAGTGAACATTCGCTATGGAGCAACAGCCCGG 3513
QY 1084 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 1103
DB 3514 GTGGTGTACAGAGAGATCGTGAGGGAGCAGCAAGGAGCCATATACAGCCTTCATC 3573
QY 1104 GluGlyLeuProGluTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123
DB 3574 GAGTCACTGCCTTAATAATATAGCACCAGATAGGAGACAAAGAACTCAGCTCTCTGGT 3633
QY 1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 1143
DB 3634 GGCAGAAACACGCAPTGGCATAGCTGTCGCTGTTAGACAGCCTCATATTTTGGCT 3693
QY 1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163
DB 3694 TTGATGAGCCACATCAGCTCTGGATACAGAAAGTGAAGAGTGTCCAGAGGCCCTG 3753
QY 1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183
DB 3754 GACAAAGCCAGAGAGGCCGCTACCTGATTTGATGCTCACCCTGTCCACCATCCAG 3813
QY 1184 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 1203
DB 3814 AATGCACACTTAATAGTGGTGTTCAGAAATGGCAGAGTCAAGGAGACGCGCACATCAG 3873
QY 1204 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAlaGlnSer 1220
DB 3874 CAGCTGTGGCAGAGGCACTATTTTCAATGTCAGTGTCCAGGCT 3924
RESULT 11
ID AAZ49333
XX AAZ49333 standard; cDNA; 3860 BP.
AC AAZ49333;
XX
DT 14-MAR-2000 (first entry)
DE Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX
KW Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW cytokine; mutant; ds.
XX
OS Synthetic.
XX Homo sapiens.
```

```
PH Key Location/Qualifiers
CDS 1..3843
FT /*tag= a
FT /product= "Human G185V mutant MDR-1 protein"
XX
PN WO9961589-A2.
XX
PD 02-DEC-1999.
XX
PE 27-MAY-1999; 99WO-US11825.
XX
PR 28-MAY-1998; 98US-0086988.
XX
PS (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PI Sorrentino B, Bunting K;
XX
DR WPI; 2000-072615/06.
DR P-PSDB; AAY58187.
XX
PT Ex vivo expansion of hematopoietic stem cells transduced with a
PT sequence encoding human multidrug resistance-1, used for bone marrow
PT transplantation -
XX
PS Example 1; Page 79-82; 113pp; English.
XX
CC This sequence represents cDNA encoding human G185V mutant multidrug
CC resistance protein MDR-1, where the Gly residue at position 185
CC of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
CC transmembrane efflux pump, responsible for the export of drugs from
CC cells, particularly cancer cells. The wild-type MDR-1 shows increased
CC resistance to etoposide and decreased resistance to vinca alkaloids
CC compared with the G185V mutant. The invention relates to transducing
CC haematopoietic stem cells with nucleic acid encoding an MDR protein
CC and culturing the modified cells. The modified haematopoietic stem
CC cells are useful in bone marrow transplantation (to reconstitute
CC haematopoietic systems in patients who have undergone chemotherapy or
CC radiation therapy) and in ex vivo gene therapy of genetic defects in
CC cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC cells can also be used to identify factors involved in regulating
CC proliferation and differentiation in haematopoietic stem cells.
CC Haematopoietic stem cells that express MDR-1 will be protected against
CC chemotherapeutic agents, so can be engrafted while the patient is
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC provides sufficient cells to permit standard biochemical analysis.
CC Overexpression of MDR-1 allows cytokine-driven expansion of
CC haematopoietic stem cells by at least 10-fold compared with a maximum
CC of 4-fold in known procedures.
XX
SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;
Alignment Scores:
Pred. No.: 0 Length: 3860
Score: 3521.50 Matches: 681
Percent Similarity: 73.92% Conservative: 246
Best Local Similarity: 54.31% Mismatches: 264
Query Match: 57.18% Indels: 63
DB: 21 Gaps: 10
US-09-873-409-5 (1-1222) x AAZ49333 (1-3860)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 151 ATGGTGGTGGAACTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 211 GTGTTTGGAAATGACAGATATCTTGGCAATGACGAAATTTAGAGATCTGATGTCA 270
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 271 AACATCACTAATAGAAGTATATCAATGATACAGGGGTCTTTCATGAATCTGGAGGAGAC 330
```



QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 331 ATGCACAGATATGCTATTATTACAGTGAATTTGGTCTGGGGTGTCTGTTCTGCTTAC 390  
QY 62 ILeGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 391 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAAGCAAAATACAAAATAGAAAACAG 450  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
DB 451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGTTGATGTCACCATGTTGGGAG 510  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 511 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTGCTGACAAAAT 570  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 571 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAGCT 630  
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 631 GGTGGAGCTAACCTTGATTTGGCCATCAGTCTCTGTTCTTGGACTGTCAGCTGCT 690  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 691 GTCGGCAAGATACACTCTTCTTACTGATAAAGAACTCTTAGCGTATGCMAAGCT 750  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 751 GGAGCAGTAGCTAGCAGAGGCTTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 810  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 811 AAGAAGAACTTGAA----- 825  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 826 -----AGGTACACAAACAAATTTAGAGAA 849  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 850 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAAATTTCTATAGTGTGCTGTTTC 909  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
DB 910 CTGCTGATCTATGCACTTATGCTCTGGCCCTCTGGTATGGGACCACCTTGGTCTCTCA 969  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 970 GGGAA-----TATTCTATTGCAAGTACTCAGTATTCTTTCTGTAATTAATGGG 1023  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 1024 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGACA 1083  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 1084 GCTTATGAAATCTTCAAGATAATTTAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1143  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 1144 GGGCACAACCATGATTAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1204 CCAATCTCGAAAAGAAAGTAAAGATCTTGAAGGGCTTGAACCTGAAGTGCAGAGTGGGACG 1263  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1264 ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGAGACAAACAGTCCAGCTGATGCAG 1323

QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
DB 1324 AGGCTCTATGACCCCCACAGAGGGGATGTCAGTGTGTGATGGACAGATATTAGGACCATA 1383  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1384 AATGTAAAGTTTCTACGGGAAATCATTTGGTGGTGTAGTCAGGAACCTGTATTGTTGCC 1443  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 1444 ACCCAGATAGCTGAAACATTCGCTATGCCGCTGAAAAATGTCCCATGTGATGAGATTGAG 1503  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1504 AAGCTGTCAAGGAAGCAATGCCCTATGACTTTATCATGAATGCCCTCATAAATTTGAC 1563  
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1564 ACCCTGGTTGGAGAGAGAGGGGCCAGTTGAGTGTGGCAGAGCAGAGGATCGCCATT 1623  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 1624 GCACGTGCCCTGGTTCCGCAACCCCAAGATCCTCTGCTGGATGAGGCCACGTCAGCCTTG 1683  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
DB 1684 GACACAGAAAGCGAAGCAGATGGTTCCAGGTGGCTCTGGATTAAGCCAGAAAAGGTGGACC 1743  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
DB 1744 ACCATTGTAGTACCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCCTGCTGTTTC 1803  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
DB 1804 GATGATGGAGTCATTGTGGAGAAAGAAATCATGTAACCTCATGAAAGAGAAAGCATT 1863  
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
DB 1864 TACTTCAAAATTTGTCAACATGCAGACAGCAGGAAATGAAGTTGAATTTAGAAAATGCAGCT 1923  
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
DB 1924 GATGAATCCAAAAGTGAATTTGATGCTTGGAAATGCTTCAATGATTCAGATCCAGT 1983  
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
DB 1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043  
QY 628 GluGluSerThrGlnSerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeu 647  
DB 2044 CTTAGTACCAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCTCTTTTGGAGGATTATG 2103  
QY 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
DB 2104 AAGCTAAATTTAACTGAATGGCCCTATTGTTGTTGGTGTATTGTCGCCATTATAAT 2163  
QY 668 GlyThrValHisProValPheSerIlePheAlaLysIleIleThrMetPheGlyAsn 687  
DB 2164 GGAGCCTCAACACCATTTGCCAATAATATTTTCAAGATATATAGGGGTTTTCACAGA 2223  
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706  
DB 2224 ATTGATGATCTCTGAAACAAACACAGAGAAATAGTAACCTGTTTCTACTATTGTTCTAGCC 2283  
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726  
DB 2284 CTTGGAATTTATTCTTTTATTACATTTTCTTCAAGGTTTTCACATTTTGGCAAGCTGA 2343  
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746  
DB 2344 GAGATCTCACCAGCGCTCCGATACATGAGTTTCCGATCCATGCTCAGACAGGATG 2403  
QY 747 AlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAsp 766



```

Db 2404 AGTTGGTTTATGACCCCTAAACACACACTGGAGCATTTGACTACCGAGGCTCGCCATGAT 2463
Qy 767 llaAaGlnleGlnGlyAlaThrGlySerArgIleGlyValleuThrGlnAsnAlaThr 786
Db 2464 GCTGCTCAAGTTAAAGGGGTATAGGTTTCAGAGCTGCTGTAATTTACCCAGATATAGCA 2523
Qy 787 AsnMetGlyLeuSerValIleIleSerPheIleIleSerPheIleIleSerPheLeuIle 806
Db 2524 AATCTTGGGACAGGATAATATATATCTTCTCATCTATGTTGGCACTAAACACTGTTACTC 2583
Qy 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db 2584 TTAGCAATTTGACCATCAATGCAATAGCAGGAGTGTTCGAAATGAAATGTTCTGGA 2643
Qy 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db 2644 CAAGCACTGAAGATAGAAGAACTAGAGAGGTCGCGGAGATCGCTACTGAAGCAATA 2703
Qy 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db 2704 GAAACCTTCGAACCGTTGTTCTTTCGACTCAGGAGCAGAGTTTGACATATGATGCT 2763
Qy 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db 2764 CAGAGTTTGCAGGTACCATACAGAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823
Qy 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db 2824 TTTTCTTCCACCCAGCAATGATATATTTTCTCATCTGATGTTTCCCGTTTGGAGCC 2883
Qy 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db 2884 TACTTGGGACATAACTCATGAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTT 2943
Qy 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db 2944 GTCCTTGGTGCCATGGCGTGGGCAAGTCAGTTTCATTTGCTGCTGACTATGCCAAAGCC 3003
Qy 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db 3004 AAAATATCAGCAGCCCATCATCATCATGATCAATGAAAAACCCCTTTGATTGACAGCTAC 3063
Qy 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db 3064 AGCAGGAAGCCCTAATGCCGAACACATTTGGAAGAAATGTCACATTTGGTGAAGTTGA 3123
Qy 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db 3124 TTCAACTATCCACCCAGCCGACATCCAGTGTCTCAGGAGCTGAGCCTGGAGGTGAAG 3183
Qy 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db 3184 AAGGCCACAGCGTGGCTCTGCTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGTGTCAG 3243
Qy 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAspAla 1046
Db 3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGAAGTCTGCTTGTATGGCAAGAAATA 3303
Qy 1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db 3304 AAGCGACTGAATGTTCACTGCTCGAGCACACCTGGGSCATCGTGTCCAGAGGCCCATC 3363
Qy 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db 3364 CTGTTTGACTCAGCATTTGCTGAGAACAATTCCTATGAGACAAACAGCCGGGTGTGTCFA 3423
Qy 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3424 CAGGAAGAGATCGTGAGGCGCAGCAAGAGGCCAACATACATGCTTCATCGATCAGTCACTG 3483
Qy 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLys 1126

```

```

Db 3484 CCTAATAATATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAA 3543
Qy 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 3544 CAACGCAATGGCCATAGCTGTCCTCTGTTAGACAGCCTCATATTTTGGTGGATGAA 3603
Qy 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db 3604 GCCACCTCAGCTCTGATACAGAAAGTGAAGAGTTGTCCAAAGACCCCTGGCAAGGCC 3663
Qy 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 3664 AGAAGAGCCGCGACCTGCAATGTGATTTGCTACCCGCTGTCCACCATCCAGATGCGAGC 3723
Qy 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 3724 TTATATAGTGGTGTTCAGATGCGCAGAGTCAGGACATGGCAGCATCAGCAGCTGCTG 3783
Qy 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 3784 GCACAGAAAGGCATCTATTTTCAATGTCAGTGTCCAGGCT 3825

RESULT 12
ABA94366
ID ABA94366 standard; DNA; 3860 BP.
XX
AC ABA94366;
XX
DT 26-MAR-2002 (first entry)
DE Human BCRP DNA related seq Id No. 3.
XX
KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatocytic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW cardiant; gene therapy; ds.
XX
OS Homo sapiens.
FH
FT Key Location/Qualifiers
FT CDS 1..3843
FT /*tag= a
XX
PN WO200192877-A2.
XX
XX 06-DEC-2001.
XX 30-MAY-2001; 2001WO-US17459.
XX 31-MAY-2000; 2000US-0584586.
XX 29-MAY-2001; 2001US-0868866.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Sorrentino B, Schuetz J;
XX
XX WPI; 2002-114368/15.
XX P-PSDB; ABB07267.
XX
XX Identifying a stem cell, for treating e.g., muscular dystrophy,
XX myocardial infarction, Parkinson's disease, or neurodegenerative
XX disorders, comprises detecting the expression of an ATP transport
XX protein (BCRP) by a cell -
XX
XX Disclosure; Page 59-60; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
XX cell that involves detecting the expression of an ATP transport protein
XX containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
XX comprising stem cells. The isolated stem cells may be used in the
XX treatment of diseases such as muscular dystrophy, degenerative liver
XX disorder, myocardial infarction, Parkinson's disease, degenerative
XX disorders of the brain, and for tissue regeneration or replacement.
XX Haematopoietic cells can be used in bone marrow transplants (e.g., for

```

CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
 CC also be used as cell targets in gene therapy protocols. The present  
 CC sequence represents a sequence related to the BCRP for which no relevant  
 CC information has been provided in the specification.

XX  
 SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 3860  
 Score: 3521.50 Matches: 681  
 Percent Similarity: 73.92% Conservative: 246  
 Best Local Similarity: 54.31% Mismatches: 264  
 Query Match: 57.18% Indels: 63  
 DB: 24 Gaps: 10

US-09-873-409-5 (1-1222) x ABA94366 (1-3860)

QY 1 MetIleuGlyIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 151 ATGGTGGTGGAACTTTGGCTGCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210  
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 DB 211 GTGTTGGAGAAATGACAGATATCTTTGCAATGCGAGAAATTTAGAAGATCTGATGTC 270  
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 DB 271 AACATCACTAATAGAGTATATCAATGATACAGGGTTCCTTCATGATCGAGGAAGAC 330  
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 DB 331 ATGACCAGATATGCTTATTATACAGTGAATTTGGTGGTGGTGGTGGTGGTGGTGGT 390  
 QY 62 IleGlyIleSerLeuThrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 DB 391 ATTCAAGTTTCATTTGGTGGCTGGAGCTGGAGAACATAACACAAATTTAGAAACAG 450  
 QY 82 PhePheIleSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
 DB 451 TTTTTCATGCTATATGCGACAGGAGATAGCTGGTTGATGTCACGATGTTGGGGAG 510  
 QY 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspLysIle 120  
 DB 511 CTTAACACCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTTGGTCAGAAAT 570  
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 DB 571 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTAC 630  
 QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 DB 631 GGTGGAGCTAACCTTTGATTTTGGCCATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 690  
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 DB 691 GTCTGGCAAGATGACTATCTTCAATCTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750  
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 DB 751 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGCTGATTCATTTGGAGACAA 810  
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220  
 DB 811 AAGAAAGACTTGA----- 825  
 QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
 DB 826 -----AGGTACAAACAAATAATTTAGAGAA 849  
 QY 241 AlaIleAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
 DB 850 GCTAAAAGAAATTTGGGATAAAGAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
 DB 910 CTGCTGATCATGTCATCTTATGCTCTGGCTTCTGTGATGGACCACTTGGTCTCTCA 969  
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 DB 970 GGGGAA-----TATTCATTTGGACAGTACTCATGTATTTCTTTCTGATTAATGGG 1023  
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
 DB 1024 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAC 1083  
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
 DB 1084 GCTTATGAAATCTTCAAGATAATTAAGCCCAAGTATTGACAGCTATTTCGAAGAGT 1143  
 QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
 DB 1144 GGGCAAAACCAAGATATATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203  
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
 DB 1204 CCATCTCGAAAGAAAGTAAAGATCTTGAAGGCGCTGAACCTGAAGTGCAGAGTGGCAG 1263  
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
 DB 1264 ACGGTGGCTGCTGGAAACAGTGGCTGTGGAGAGACACACAGCTCCAGCTGATGTCAG 1323  
 QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
 DB 1324 AGGCTCTATGACCCACAGAGGGGATGGTCAGTTGTGATGGACAGGATATTAGGACCAT 1383  
 QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
 DB 1384 NATGTAAGTTTCTACGGGAATCATTTGTGTGTGAGTCAGGAACTGTTATTTGTTGCC 1443  
 QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460  
 DB 1444 ACCACGATAGCTGAAACATTCGCTATGCGCGTGAATAATGTCCACCATGTGATGAGATTG 1503  
 QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
 DB 1504 AAAGCTGTCAAGGAAGCCCAATGCCATATGATCTTATCATGAAACTGCTCTATAAATTG 1563  
 QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
 DB 1564 ACCCTGTTGGAGAGAGAGGCGGCCAGTTGAGTGTGGCGAGAGAGAGGATCGCCATT 1623  
 QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
 DB 1624 GCACGTGCTGCTGTTGCAACCCCAAGATCTCTCTGATGGAGGCCAGCTCAGCCTTG 1683  
 QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
 DB 1684 GACACAGAAGCGAGCAGTGTTCAGGTGGTCTCTGGATAAGGCCAGAAAGTCGGACC 1743  
 QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
 DB 1744 ACATTTGATGATGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGTTTC 1803  
 QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
 DB 1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAATCATATGAAGAGAAAGCAT 1863  
 QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
 DB 1864 TACTTCAAACTGTCTCAATTCAGAGAGAGGAAATGAAGTTGAATTAGAAAATGCGAGCT 1923  
 QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
 DB 1924 GATGAATCCAAAGTGAAATGATGCTTGGAAATGCTCTTCAAAATGATTCAGATCCAGT 1983



XX  
PR 04-NOV-1999; 99US-0163508.

PA (INCY-) INCYTE GENOMICS INC.

PI Sornasse T, Seilhamer JJ, Watson GA;

WPI: 2001-291057/30.

PT New cell and tissue specific polynucleotides useful for diagnosis, is  
PT prognosis or monitoring of treatments for disorders where the gene is  
PT associated with a cancer, immunopathology or neuropathology -

PS Claim 1; Page 207-208; 327pp; English.

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide sequences (I). (I) can have cytostatic, immunomodulatory and neuroprotective activities, and can be used in gene therapy. (I) and proteins (II) encoded by then are used in high throughput screening assays to select DNA molecules, RNA molecules, peptide nucleic acids, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents. Expression of (I) in a sample indicates the differentiation of embryonic stem cells into a tissue selected from brain, heart, kidney, liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used to produce an expression profile that defines a metabolic or developmental process, treatment, condition, disease or disorder. The gene profile can be used for diagnosis, prognosis or monitoring of treatments and for investigating a predisposition to a disorder where the gene is associated with a cancer, immunopathology or neuropathology.

Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

**Alignment Scores:**

Pred. NO.:	0	Length:	4349
Alignment Score:	3521.50	Matches:	681
Score:		Conservative:	246
Percent Similarity:	73.9%	Mismatches:	264
Best Local Similarity:	54.31%	Indels:	63
Query Match:	57.18%	Gaps:	10
DB:	22		

US-09-873-409-5 (1-1222) X AAH57442 (1-4349)

Qy	1	Met	Ile	Leu	Gly	Ile	Leu	Ala	Ser	Leu	Val	Asn	Gly	Ala	Cys	Leu	Pro	Leu	Met	Pro	Leu	20
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	276	ATGT	GGT	GGG	AA	CTTT	GG	TGC	CAT	CAT	CC	ATG	GGG	GTG	GAC	TTT	CCT	CT	CAT	GTG	CTG	335

```

Qy      21  valLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
          :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:

```

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43

Db 396 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCCTCATGAATCTGGAGGAAGAC 455

456 ATGACCAGGATGCGCTATTATTACAGTGAATTGGTCTGGGGGTCTGGTTGCTGCTTAC 515

Qy 62 IleGlnIleSerLeuTrrpileIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81

QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101

Db 576 TTTTTCATGCTATAATGCCACAGGAGATAGGCTGGTTTCATGTGCACGATGTTGGGAG 635

Db

636 CTTAACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTGGTGACAAAAATT 695

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140

Db	696	GGAAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGTATTATAGTAGGATTTACAGCT	755
Qy	141	GlyTriPlysLeuThrLeuValThrLeuSerThrSerProLeuMetAlaSerAlaAla	160
Db	756	GGTTGGAAGCTAACCCCTTGATGATTTTGGCCATCAGTCTCTGTTCTGGAGTGTGCACTGCT	815
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	816	GTCGTGGCCAAAGATATCTATCTCTTACTGATAAAGAACTCTTACGCGTATGCCAAAGCT	875
Qy	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	876	GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGACCTGTGATTGCATTTTGGAGGACAA	935
Qy	201	GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe	220
Db	936	AAGAAAGAACTTGAA-----	950
Qy	221	ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp	240
Db	951	-----AGGTACAAACAAAATTTTGAAGAA	974
Qy	241	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	260
Db	975	GCTAAAGAAATTGGGATAAAGAAAGCTATTACAGCCATATTTCTATAGGTGCTGCTTTC	1034
Qy	261	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsn	280
Db	1035	CTGCTGATCTCATCTTATGCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCTCA	1094
Qy	281	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	300
Db	1095	GGGGAA-----TATTCTATTGGACAAGTACTCCTGTATTTCTTTCTGTATTAATTCGG	1148
Qy	301	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	320
Db	1149	GCCTTTAGTGTGGACAGGCATCTCCAGCATTTGAAGCATTTGCCAATTCAGAGGAGCA	1208
Qy	321	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	340
Db	1209	GCCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGCACAGCTATTTCGAAGAGT	1266
Qy	341	GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr	360
Db	1269	GGGCACAAACACGATAATATTAGGGAAATTTGGAAATTCACAAATGTTTCACCTTCAGTTAC	1328
Qy	361	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	380
Db	1329	CCATCTCGAAAAGAAGTTAAGATCTTTGAAGGGCCTGAACCTGAAGGTGCAGASTGGGCAG	1388
Qy	381	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	400
Db	1389	ACGCTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAAACAGTCGAGCTGATGCAG	1448
Qy	401	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	420
Db	1449	AGGCTCTATCACCCACAGAGGGGATGGCTGAGTTGTGATGGACAGGATATTAGGACCATA	1508
Qy	421	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	440
Db	1509	AATGTAAAGTTTCTACGGGAAATCATTTGGTGTGGTGAAGTGTATGTGTTTTCCTTC	1568
Qy	441	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluLeuMetGlu	460
Db	1569	ACCACGATAGCTGAAACATTCCTGATGGCCGTGAAATGTCACCATGGATGAGATTGAG	1628
Qy	461	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	480
Db	1629	AAAGCTGTCAAGGAAGCAATGCCTATGACTTTTATCATGCAATCGCTCATAAATTTGAC	1688
Qy	481	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	500



QY 1207 ArgAsnArgAspIleTyrPheIysLeuValAsnAlaGlnSer 1220  
 Db 3909 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 3950

## RESULT 14

AAZ94738  
 ID AAZ94738 standard; cDNA; 4646 BP.

AC AAZ94738;

DT 01-AUG-2000 (first entry)

DE Human ATP binding cassette ABCB1 (MDR1) cDNA.

KW ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;  
 KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;  
 KW lupus erythematosus; diagnosis; gene therapy; MDR1;  
 KW multidrug resistance; chromosome 7q21; ss.

XX Homo sapiens.

XX WO200018912-A2.

XX 06-APR-2000.

PF 21-SEP-1999; 99WO-EP06991.

PR 25-SEP-1998; 98US-0101706.

PA (FARB ) BAYER AG.

XX Schmitz G, Klucken J;

DR WPI; 2000-293151/25.

XX Adenosine triphosphate binding proteins useful for identifying agents  
 PT for treating atherosclerosis and other inflammatory disorders -

PS Claim 9; Page 110-112; 154pp; English.

XX The present sequence is that of human ATP binding cassette  
 CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a  
 CC differential display method in which monocytes from peripheral  
 CC blood were subjected to macrophage differentiation and cholesterol  
 CC loading with acetylated low density lipoproteins and subsequent  
 CC cleaving with high density lipoprotein (HDL3) to identify  
 CC cholesterol sensitive genes. The gene maps to chromosome 7q21  
 CC and is also termed MDR1 (multidrug resistance). The invention  
 CC provides cholesterol-sensitive ABC genes (see AAZ94734-63). These  
 CC genes, and polypeptides encoded by them, can be used for diagnostic  
 CC and therapeutic applications, and for biochemical or cell-based  
 CC assays to screen for pharmacologically active modulator compounds  
 CC useful for the treatment of lipid disorders, atherosclerosis or  
 CC other inflammatory diseases such as psoriasis and lupus  
 CC erythematosus.

XX SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 4646  
 Score: 3521.50 Matches: 681  
 Percent Similarity: 73.92% Conservative: 246  
 Best Local Similarity: 54.31% Mismatches: 264  
 Query Match: 57.18% Indels: 63  
 DB: 21 Gaps: 10

US-09-873-409-5 (1-1222) x AAZ94738 (1-4646)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 575 ATGGTGGTGGGAACCTTTGGCTGCATCATCCATGGGGCTGGACTCTCTCATGATGCTG 634

QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 635 GTGTTGGAGAAATGACAGATATCTTTGCAATGAGGAAATTTAGAGATCTGATGCA 694  
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 695 AACATCAGCTAATAGAAGTGATATCAATGATACAGGGTTCTTTCATGAATCGGAGGAGAC 754  
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 755 ATGACCAGGTATGCCTATTATTATACAGTGAATTTGGTGGGGTGTGCTGCTTAC 814  
 QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 815 ATTCAGGTTTCATTTTGGGCTCGCAGCTGGAGCTGGAACAAATACACAAATTTAGAAACAG 874  
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
 Db 875 TTTTTCATGCTATAATGCGACAGGAGATAGCGTGGTTGATGTGCACGATGTGGGAG 934  
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 935 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTGGTGACAAATTT 994  
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 995 GGAATGTTCTTTCAGTCAATGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAGCT 1054  
 QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 1055 GGTGGAGCTAACCTTGATTTTGGCCATCAGTCCCTGTTCTTGGACTGTGACGTGCT 1114  
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 1115 GTCTGGCAAGATATCTATCTTCACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174  
 QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 1175 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATTCATTTGGAGACAA 1234  
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
 Db 1235 AGAAAGAACTTGAA----- 1249  
 QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
 Db 1250 -----AGGTACAACAAAAAATTTAGAGAA 1273  
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
 Db 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATTTTCTATAGGTGCTGCTTC 1333  
 QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
 Db 1334 CTGCTCATGTCATCTTATGCTCTGGCTTCTGGTAIGGGACCACTTGGTCTCTCA 1393  
 QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 Db 1394 GGGGAA-----TATTCTATTGGCAAGTACTCAGCTGATTTCTTTCTGATTAAATTGGG 1447  
 QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
 Db 1448 GCTTTTAGTGTGGACAGCATCTCCAAGCATTTGAAGCATTTTGCAAATCGACAGGAGCA 1507  
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
 Db 1508 GCTTATGAATCTTCAAGATAATTGATAATAGCCCAAGTATTGACAGCTATTCCGAAGAGT 1567  
 QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
 Db 1568 GGGCACAACACAGATAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627  
 QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380



Db 1628 CCATCTCGAAGAGAGTAAGATCTTGAAGGGCCCTGAACTCGAAGGTGCAGAGTGGGCAG 1687  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1688 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGAGAGACAAACAGTCCAGCTGATGCAG 1747  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1748 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGCAGCAGGATATTAGGACCAT 1807  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1808 AATGTAAGGTTTCTACGGGAATCAATGGTGTGTGAGTCAGGAACCTGATTGTTGGC 1867  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1868 ACCACGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAG 1927  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1928 AAGCTCTCAAGAGAGCCAAATGCTATGACTTATCATGAACCTGCCTCATAAATTGAC 1987  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1988 ACCCTGGTTGGAGAGAGGGCCAGTTCAGTGTGTGGCAGAGCAGAGGATCGCCATT 2047  
QY 501 AlaArgAlaValArgAsnProLysIleIleuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 2048 GCAGTGCCTGGTTCGCAACCCCAAGATCCTCTGCTGGATGAGGCACGTCAGCCCTG 2107  
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 2108 GACACAGAGAGGAGCAGTGGTTCAGTGGCTCTGGTAAGGCCACAAAGGTCCGACC 2167  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 2168 ACCATTGTGATGATCATCGTTTCTACAGTTCGTAATGCTGACGTCGCTCGTGGTTTC 2227  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
Db 2228 GATGATGGAGTCATGTGGAGAGAGGAATCATGATGAATCATGAAGAGAGAGGCAATT 2287  
QY 581 TyrTyrSerLeuLeuMetSerGln-----AspIleLysLysAlaAsp 594  
Db 2288 TACTTCAAACTGTTCACAAATCCACACAGCAGGAATGAAGTTGAATTAGAAATGCAGCT 2347  
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 2348 GATCAATCCAAAGTGAAATGTAGTCCCTGGAAATGTCTTCAAAATGATTCAGATCCAGT 2407  
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
Db 2408 CTAATAAGAAAGAGATCAACTCGTAGGAGTCTCGTGGATCACAAGCCCAAGACAGAAAG 2467  
QY 628 GluGluSerThrGlnSerLysLeuIleSerLeuProGluValSerLeuLysIleLeu 647  
Db 2468 CTTAGTACCAAGAGGCTCGATCAAGATATACCTCAGTTTCTTTGGAGGATTATG 2527  
QY 648 LysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
Db 2528 AAGCTAAATTTAACTGAATGCCCTATTATTTGTTGGTGTATTGTCGCAATTATAAAT 2587  
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687  
Db 2588 GGAGGCTGCAACACGAGATTTGCAATAATATTTCAAGAATATATAGGGGTTTTTACAGA 2647  
QY 688 ---AsnAspLysThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706  
Db 2648 ATTGATGATCTGAAACAAACAGCAGATAGTACTGTTTTCATTACTATTCTTCTAGCC 2707  
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726  
Db 2708 CTTGGAATTAATTTCTTTTATACATTTTCTTTCAGGGGTTTTTCAATTTTGGCAAGCTGGA 2767

QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746  
Db 2768 GAGATCCTCACCAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTG 2827  
QY 747 AlaTyrPheAspGluLysGluAsnSerThrGlyLysLeuThrThrIleLeuAlaIleAsp 766  
Db 2828 AGTGTGTTTATGATCCCTAAACACACCTGGAGCATTTGACTACCAAGGCTCGCCAATGAT 2887  
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786  
Db 2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTTCAGGCTTCTGCTAATTTACCCAGATATAGCA 2947  
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIle 806  
Db 2948 AATCTTGGGACAGGAATAATTATATCTTCATCTATGTTGGCACTAACACTGTTACTC 3007  
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826  
Db 3008 TTAGCAATTGTACCCCATCATTCATATGACGAGGAGTTGTTGAAATGAAATGTTGCTGGA 3067  
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846  
Db 3068 CAAGCACTGAAGATAGAAAGAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127  
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866  
Db 3128 GAAACTTCGGAACCGTGTCTTTGACTCAGGAGCAGAAGTTTGAACATATGATGCT 3187  
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886  
Db 3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGACACACATCTTTGAAATTACA 3247  
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906  
Db 3248 TTTTCTTCCACCAGGCAATGATATTTTCTATGCTGGATGTTTCCGGTTTGGAGCC 3307  
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926  
Db 3308 TACTTGGTGCACATAAACTCATGAGCTTTGAGGATGTTCTCTAGTATTATTTTCAGCTGTT 3367  
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946  
Db 3368 GTCTTTGGTGCCTATGCGGTGGGCAAGTCAGTTTCATTTCTCTGCTGGATGTTTCCGGTTTGGAGCC 3427  
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966  
Db 3428 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAGAAATGTCACATTTGGTGAAGTTGTA 3487  
QY 967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986  
Db 3488 AGCAGGAAGCCCTAATGCCGAACACATTCGAAGGAATGTCACATTTGGTGAAGTTGTA 3547  
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006  
Db 3548 TTAACATATCCACCCGACCGGACATCCAGTGTCTCAGGAGCTGAGCCCTGGAGGTGAAG 3607  
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026  
Db 3608 AAGGGCCAGACGCTGGCTCTGGTGGGCGAGGTGGTGGGAGAGACACAGTGGTCCAG 3667  
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAla 1046  
Db 3668 CTCTGGAGCGGTTCTACAGCCCTTGGCAGGGAAGTGTGCTTGTATGCAAGAAATA 3727  
QY 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066  
Db 3728 AAGCGATTAATGTTCACTGGCTCCGAGCACACCTCGGCACTCGTGTCCAGGAGCCCAT 3787  
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086  
Db 3788 CTGTTTGACTGACGATTCCTTGAGAAATTCCTGATGAGACAAACCCGGGTGTGTCA 3847



```
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db 3848 CAGGAAGAGATCGTAGGGCAGCAAGAGGAGGCCAACATACATCGCTTCATCGAGTCACTG 3907
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db 3908 CCTAATAATATAGCACTAAAGTAGGAGCAAGAGAACTCAGCTCTCTGGTGCCAGAAA 3967
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db 3968 CAACGCATTGCCATAGCTCGTGCCTTGTAGACAGCCTCATATTTGCTTTGGATGAA 4027
QY 1147 AlaThrSerAlaLeuAspAsnSerGluLysValGlnHisAlaLeuAspLysAla 1166
Db 4028 GCCACGCTCAGCTGATGATACAGAAAGTGAAGGTTGTCCAAAGACCTCGGCAAGCC 4087
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db 4088 AGAAGAGCGCCACCTGCATTGTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGAC 4147
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db 4148 TTAATAGTGGTGTTCAGATGCGCAGAGTCAGAGCATGGCAGCATCAGCAGCTGCTG 4207
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

RESULT 15
AAD38994
ID AAD38994 standard; DNA; 4646 BP.
XX
AC AAD38994;
DT 23-SEP-2002 (first entry)
XX
DE Human mdrl gene.
XX
KW Human; hematologic malignancy; multidrug resistance; MDR; SUMO-1;
KW hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-1;
KW lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
KW myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
KW angogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
KW polycythaemia vera; hypoxia responsive element; HRE; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 425..4267
FT FT /*tag= a
FT FT /product= "Human MDR protein"
XX
PN WO200234291-A2.
XX
PD 02-MAY-2002..
XX
PF 25-OCT-2001; 2001WO-US49856.
XX
PR 26-OCT-2000; 2000US-243542P.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Colgan SP;
XX
PI WPI; 2002-471427/50.
DR P-PSDB; AAE24211.
XX
PT Treating a subject (at risk of) having a hematologic malignancy or
PT multidrug resistance, e.g. lymphoma or myeloma, by administering
PT hypoxia inducible factor 1 binding molecules or small
PT ubiquitin-like-modifier-1 binding molecules -
XX
PS Disclosure; Page 53-55; 92pp; English.
```

```
XX The invention relates to a method of treating a subject having or at
CC risk of developing a haematologic malignancy or multidrug resistance
CC (MDR). The method involves administering hypoxia inducible factor-1
CC (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1
CC binding molecules or HIF-1-SUMO-1 complex modulators. mdrl-hypoxia
CC responsive element (HRE) binding molecules or antisense nucleic
CC acid molecules and SUMO-1 binding molecules or antisense molecules
CC are useful for treating a subject having or at risk of developing
CC haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid
CC disorder). The lymphoid disorders include lymphocytic leukaemia or
CC chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic
CC lymphoid leukaemia. The myeloid disorders include chronic or acute
CC myeloid leukaemia, e.g. angioleukemia myeloid metaplasia, essential
CC thrombocythaemia or polycythaemia vera. The invention is used in gene
CC therapy. The present sequence is human mdrl gene.
XX
SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4646
Score: 3521.50 Matches: 681
Percent Similarity: 73.92% Conservative: 246
Best Local Similarity: 54.31% Mismatches: 264
Query Match: 57.18% Indels: 63
DB: 24 Gaps: 10

US-09-873-409-5 (1-1222) x AAD38994 (1-4646)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACCTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAAATGACATATCTTGCAAATGCAGGAATTTAGACATCTGATGTCTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAAGAGTGATATCAATGATACAGGGTCTTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACACCGATATGCTCTATTATACAGTGGAAATTTGGTGGGGTGTGCTGCTGCTTAC 814
QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCATTTTGGTGCTGGCAGCTGGAAGCAAAATACACAAAATTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTAATAATGCAGCAGGAGATAGGCTGGTTTGTGTCACCATGTTGGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAATGAAGTATTATGGTGCAAAATT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAAGCTAACCTTGTGATTGTTGGCCATCAGTCTCTGTTCTTGAGCTGTCAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACATCTCTCAATTTACTGATAAAGAACTCTTAGCTGATGCAAAAGCT 1174
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGCTCTTGGCAGCAATTAGAACTGTGATTGCTTGGAGGACAA 1234
```

QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106

Db 3848 CAGGAAGAGATCGTAGGGCAGCAAGAGGAGGCCAACATACATCGCTTCATCGAGTCACTG 3907

QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126

Db 3908 CCTAATAATATAGCACTAAAGTAGGAGCAAGAGAACTCAGCTCTCTGGTGCCAGAAA 3967

QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146

Db 3968 CAACGCATTGCCATAGCTCGTGCCTTGTAGACAGCCTCATATTTGCTTTGGATGAA 4027

QY 1147 AlaThrSerAlaLeuAspAsnSerGluLysValGlnHisAlaLeuAspLysAla 1166

Db 4028 GCCACGCTCAGCTGATGATACAGAAAGTGAAGGTTGTCCAAAGACCTCGGCAAGCC 4087

QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186

Db 4088 AGAAGAGCGCCACCTGCATTGTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGAC 4147

QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206

Db 4148 TTAATAGTGGTGTTCAGATGCGCAGAGTCAGAGCATGGCAGCATCAGCAGCTGCTG 4207

QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220

Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 1235 AAGAAAGAACTTGA 1249  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLeuAsp 240  
Db 1250 -----AGGTACAACAAAAATTTAGAGAA 1273  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 1274 GCTAAAGAAATTTGGGATAAGAAAGCTATTACAGCCAATATTCTATAGGTGCTGCTTC 1333  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrPheTyrGlyThrSerLeuIleLeuAsn 280  
Db 1334 CTGCTGATCTATGCTTATGCTCTGGCTTCTGGTATGGGACCCTTGGCTCTCA 1393  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1394 GGGGAA-----TATTCTATTGGACAAGTACTCAGTGTATTCTTTCTGTATTAAATTGGG 1447  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1448 GCCTTTAGTGTGGACAGGCACTCCCAAGCATTTGAACATTTGCAATGCAAGAGAGCA 1507  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1508 GCCTTATGAATCTTCAAGATAATTAATAAGCAAGTATTGACAGTATTGCAAGAGT 1567  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1568 GGGCACAAACACAGATAATTAAGGGAAATTTGCAATTCAGAAATGTTCACTTCAGTTAC 1627  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1628 CCATCTCGAAAGAAAGTAAAGTCTTGAAGGGCTGCAACCTGAAGTGCAGAGTGGGCA 1687  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1688 ACGTGGCCCTGGTGGAAACAGTGGCTGTGGGAGAGCAACAGTCCAGCTGATGCGAG 1747  
QY 401 ArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleAlaLeu 420  
Db 1748 AGGCTCTATGACCCACAGAGGGAGTGGTCAGTGTGATGACAGGATATTAGGACCAT 1807  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1808 AATGTAAAGTTTCTACGGGAAATCAITGGTGTGAGTGAAGCAACCTGATTGTTGCC 1867  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460  
Db 1868 ACCACGATAGCTGAAACATTCGCTATGGCCGTGCAAAATGTCACCATGGATGAGATTGAG 1927  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1928 AAAGCTGTCAAGGAAGCAATGCTATGACTTTTATCATGAACCTGCTCATAAATTTGAC 1987  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1988 ACCCTGTTGGAGAGAGAGGGGCCAGTTGAGTGGTGGGAGAGAGAGATGCGCATTT 2047  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 2048 GCAGTGCCTTGGTTCGCAACCCCAAGATCTCTGCTGATGAGGCCAGCTCAGCCTTG 2107  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysLysAlaSerLysGlyArgThr 540  
Db 2108 GACACAGAAAGCGAAGCAGTGGTTTCAAGTGGCTCTGGATAAGGCCAGAAAGGTCGACC 2167  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 2168 ACCATTGTGATGATCATCTGTTGCTACAGTTCGTAAGTTCGTAAGTTCGCTGCTGTTTC 2227  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580

Db 2228 GATGATGGAGTCATTCTGGAGAAAGGAAATCATGATGAACATGAAAGAGAAAGGCATT 2287  
QY 581 TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp 594  
Db 2288 TACTTCAAACTTTGTCAATGACAGACAGCAGGAAATGAAGTTTGAATAGAAAATGCGACT 2347  
QY 595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614  
Db 2348 GATGAATCCAAAAGTGAAATGATGCTTGGAAATGCTCTCAATGATTCAAGATCCAGT 2407  
QY 615 SerVal-----LysSerIleLys---SerAspPheIleAspLysAla 627  
Db 2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAAGCCACAGAGAA 2467  
QY 628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647  
Db 2468 CTTAGTACCAAAAGAGCTCTGGATGAAGTATACCTCCAGTTTCTTTTGGAGGATATG 2527  
QY 648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667  
Db 2528 AGCTAAATTTAACTGAATGGCCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2587  
QY 668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687  
Db 2588 GGAGGCTCGAACCCAGCATTTTGAATTAATTTTCAAGATTTATAGGGTTTTTCACAGA 2647  
QY 688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706  
Db 2648 ATTGATGATCTCTGAAACAAACACGACAGAATAGTAACCTTGTTTTCACTATTGTTCTAGCC 2707  
QY 707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726  
Db 2708 CTTGGAAATTTCTTTTATTAATTTTCTTTCAGGGTTTCACTTTGGCAAGCTGGA 2767  
QY 727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746  
Db 2768 GAGATCCTCACCAAGCGCTCCGATACATGGTTTCCGATCCATGCTCAGACAGGATGTG 2827  
QY 747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIleLeuAlaIleAsp 766  
Db 2828 AGTGTGTTTGTATGACCTTAAACACCACTGGAGCAATTGACTACCAAGGCTGCGCAATGAT 2887  
QY 767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786  
Db 2888 GCTGCTCAAGTTAAAGGGCTATAGTTCCAGGCTTGTGTAATTACCCAGAAATATAGCA 2947  
QY 787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806  
Db 2948 AATCTTGGACAGGAATAATTATATCTTCACTATGTTGGCAACTAACACTGTACTC 3007  
QY 807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826  
Db 3008 TTAGCAATTTGTACCCATCATTCGAATGAGGAGGTTGTTGAAATGAAATGTTGTCTGGA 3067  
QY 827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846  
Db 3068 CAGGACTGAAAGATGAAGAACTAGAGGGTGTGGGAAGATCGCTACTGAGCAATA 3127  
QY 847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866  
Db 3128 GAAAACTTCGGAACCGTTGTTCTTGTACTCAGGACAGAGTTTCAACATATGTTATGCT 3187  
QY 867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886  
Db 3188 CAGAGTTTTCAGGTACCATACAGAAATCTTTTGGAGAAAGCACACATCTTTGGAATTACA 3247  
QY 887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906  
Db 3248 TTTTCTTCCACCGAGCAATGATGATTTTCTTCTGCTGGATGTTTCCGTTTGAGGCC 3307  
QY 907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926

Db 3308 TACTTGGTGGCACATAAATCATAGACCTTTGAGGATGTTCTGTAGTATTTCAGCTGTT 3367  
QY 927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946  
Db 3368 GTCTTTGGTGCCATGGCCCTGGGCAAGTCAGTTTCATTTCCTGACTATGCCAAGCC 3427  
QY 947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysPheProAsnIleAspSerArg 966  
Db 3428 AAAATATCAGCAGCCACATCATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTAC 3487  
QY 967 SerGlnGluGlyLysProAspThrCysGluGlyAsnLeuLysPheArgGluValSer 986  
Db 3488 AGCAGGAAGGCTTAATGCCGAACATTTGGAAGGAAATGTACATTTGGTGAAGTTGTA 3547  
QY 987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006  
Db 3548 TTCNACTATCCACCCAGCCGACATCCCAAGTCTTCAGGGACTGAGCCTGGAGGTGAAG 3607  
QY 1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026  
Db 3608 AAGGGCCAGACGCTGCTGTGGTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667  
QY 1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046  
Db 3668 CTCCTGGAGCGGTTCTACGACCCCTTTGGCAGGGAAGTGTCTTGTATGGCAAGAAATA 3727  
QY 1047 LysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066  
Db 3728 AAGCGACTGAATGTTTCAGTGGCTCCAGACACACCTGGGCATCGTGTCCAGGAGCCCATC 3787  
QY 1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086  
Db 3788 CTGTTTGACTGCAGCATTTGCTGAGAACATTTGCCATATGGAGACAACAGCCGGTGGTGTC 3847  
QY 1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106  
Db 3848 CAGGAAGAGATCGTAGGGCAGCAAGAGGAGCCCAACATACATGCCCTTCATCGAGTCAC 3907  
QY 1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyValAlaGlnLeuSerGlyGlnLys 1126  
Db 3908 CCTAATAATATAGCACTAAAGTAGGAGACAAGGAACCTCAGCTCTCTGGTGCCAGAAA 3967  
QY 1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGlu 1146  
Db 3968 CRAACGATTGCCATAGCTCGTGCCCTTTGTAGACAGCCTCATATTTTGTGATGAA 4027  
QY 1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166  
Db 4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAGAAAGCCCTGGACAAAGCC 4087  
QY 1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186  
Db 4088 AGAAGAGCCGCACTGCTGATTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGAC 4147  
QY 1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206  
Db 4148 TTAATAGTGGTGTTCAGAAATGCGAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTG 4207  
QY 1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 4208 GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

Search completed: March 30, 2003, 03:16:38  
Job time : 1426.3 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 16520.1 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-5  
Perfect score: 6159  
Sequence: 1 MILGILASLVNGACPLMPL.....QELLENRDIYFKLVNAQSVQ 1222

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USFTO.spool/US09873409/runat\_27032003.115418.19228/app\_query.fasta\_1.7544  
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 54887 @runat\_27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLIPF -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_tgo\_hum.\*  
40: em\_tgo\_mus.\*  
41: em\_tgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	6157	100.0	3702	6 AX339031	AX339031 Sequence
2	5988.5	97.2	3621	6 AX339032	AX339032 Sequence
3	5737	93.1	3699	6 AX478104	AX478104 Sequence
4	5328	86.5	3177	6 AX339030	AX339030 Sequence
5	4561.5	74.1	2856	6 AX339028	AX339028 Sequence
6	3550.5	57.6	4296	10 CRUPGPI	M60040 C.griseus P
7	3549.5	57.6	4304	10 CRUPGPI165	M59253 Chinese ham
8	3533.5	57.4	4279	6 AX105082	AX105082 Sequence
9	3533	57.4	4195	6 AX108656	AX108656 Sequence
10	3532	57.3	4927	10 AF257746	AF257746 Rattus no
11	3530.5	57.3	3860	6 AX322787	AX322787 Sequence
12	3530.5	57.3	4279	6 AX105078	AX105078 Sequence
13	3530.5	57.3	4279	6 AX105080	AX105080 Sequence
14	3530.5	57.3	4646	6 AX391099	AX391099 Sequence
15	3528.5	57.3	4045	12 AF269224	AF269224 Synthetic
16	3527.5	57.3	4378	6 E02326	E02326 Multidrug r
17	3526	57.2	3858	4 OAU78609	OAU78609 Ovis aries
18	3525.5	57.2	4279	6 AX105057	AX105057 Sequence
19	3525.5	57.2	4317	4 AF045016	AF045016 Canis fam
20	3525.5	57.2	4317	6 AX105059	AX105059 Sequence
21	3524.5	57.2	4186	6 AX108654	AX108654 Sequence
22	3521.5	57.2	3860	6 AX332789	AX332789 Sequence
23	3521.5	57.2	4646	6 AX336420	AX336420 Sequence
24	3521.5	57.2	4646	6 AX336708	AX336708 Sequence
25	3521.5	57.2	4646	6 I49610	I49610 Sequence 2
26	3521.5	57.2	4646	9 HUMMDR1	M14758 Homo sapien
27	3521.5	57.2	4669	6 I08557	I08557 Sequence 3
28	3521.5	57.2	6505	6 AR028671	AR028671 Sequence
29	3521.5	57.2	8630	6 AX012320	AX012320 Sequence
30	3521.5	57.2	8630	6 AX012321	AX012321 Sequence
31	3521.5	57.2	9318	6 AR028672	AR028672 Sequence
32	3520	57.2	4264	6 AR051647	AR051647 Sequence
33	3520	57.2	4264	6 AR051650	AR051650 Sequence
34	3518	57.1	4323	10 AF286167	AF286167 Rattus no
35	3514	57.1	3840	6 AX481416	AX481416 Sequence
36	3514	57.1	3988	6 AX024454	AX024454 Sequence
37	3514	57.1	4192	9 AF016535	AF016535 Homo sapi
38	3505.5	56.9	4280	10 CRUPGPII	M60041 C.griseus P
39	3503.5	56.9	4669	6 AR091275	AR091275 Sequence
40	3503.5	56.9	4669	6 AR203322	AR203322 Sequence
41	3503	56.9	4189	6 AX322791	AX322791 Sequence
42	3503	56.9	4298	10 MUSMDR	M14757 Mouse multi
43	3499.5	56.8	3834	4 CFA419568	AJ419568 Canis fam
44	3499	56.8	3888	10 AY082609	AY082609 Rattus no
45	3499	56.8	4788	6 AX322793	AX322793 Sequence

ALIGNMENTS

RESULT 1

AX339031  
LOCUS AX339031 3702 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 13 from Patent WO0194400.  
ACCESSION AX339031  
VERSION AX339031.1 GI:18129123  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
JOURNAL homologue on chromosome 7p15-21 and uses thereof  
Patent: WO 0194400-A 13 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source  
1. 3702  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3702  
Score: 6157.00 Matches: 1222  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.97% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-5 (1-1222) x AX339031 (1-3702)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 34 ATGATCTCGGTATACCTGGCATCTGGTCAATGGAGCCCTGCCTTCCTTAATGCCACTG 93  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 94 GTTTTAGAGAAATGAGTGATTAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153  
Qy 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
Db 154 TCCTTCTCAGGTTGACCCGTATTATGTTGGAATAGGTGTGCTGCTTCGATTTTGGT 213  
Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
Db 214 TACATACAGATTTCTCTGTGGATTTAATCTGACACGACACGACGAGAGATTCGAAAA 273  
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
Db 274 CAGTTTTTTTCATTGCTTTGGCAGCAGACATCGGCTGGTTGTATAGCTGTGACATCGGT 333  
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 334 GAACCTTAACACTCGCATCGACAGACATTGACAAATCAGTGATGTTATGGAGATAAGATT 393  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCGATTGGCTGGCAGTGGTTGGTGAAG 453  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 454 GGCTGGAAACTCACCTTAGTACCTATTCACAGTCTCTCTTATTAATGGCTTCAGCGGCA 513  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 514 GCATGTTCTAGGATGGTCATCTCATTCACAGTAAGGAATTAAGTGCTTATTCCAAAGCT 573  
Qy 181 GlyAlaValAlaGluValLeuLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 574 GGGGCTGTGGCAGAGAAGTCTTGTCATCATCCGAAACAGTCATAGCCTTTTAGGGCCAC 633

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 634 GAGAAAGAACTTCAAAGGTCTTTCCCTTTTAAATATAACAAGATATGCTTGGTTTATTTT 693  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 694 CCCAGTGGCTACTAAGTTGTGTCTGTTNTTGTAAAGGTATACACAGAAATCAAGAT 753  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 754 GCAAAGATTTTGGCATAAAGAGCTATAGCTTCAAAGTGCTCTCTGGTCTGTGTATC 813  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
Db 814 TTCTTTATGAATGGAACCTATGACCTTGTCTTTTGGTATGGAACCTCTCTTGAATCTTA 873  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 874 GGAGAACTCGATATACCATCGGAGCTGTCTTCTGCTGTTTCTTAGTGTAAATCCATAG 933  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 934 AGTTATTGCATTGGAGCAGACGTCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCT 993  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 994 GCCTTTTCATATTTCCAGGTTATTGATAAGAAACCCAGATATAGATAACTTTCCACAGCT 1053  
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1054 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAAAATGTTCTTCAATFAT 1113  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1114 CCATCAAGACCATCTATCAGATTCTGAAAGGTCTGAACTCAGAAATTAAGTCTGGAGAG 1173  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1174 ACAGTCGCTTGTGCTGCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTCGAG 1233  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1234 AGGTTATATGATCCGATGATGCTTATCATGCTGGATGAGAATGACATCAGAGCTTTA 1293  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1294 AATGTCGGGCTTATCGAGACCATTTGGAGTGTGTAGTCAAGAGCCTGTTTGTTCGGG 1353  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATCGAG 1413  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1414 AGAGCAGCAAGGAAGCAATGGCTATGATTTTATCATGGAGTTTCCATAAATTTAAT 1473  
Qy 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1474 ACATTGTGGGAAAGAGAGCTCAATAGTGGAGGCGAGAACACAGAGATCGCANT 1533  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1534 GCTCGTGCTTAGTTCGAAACCCCAAGATTCTCATTTTAGATGAGGCTAGCTCTGCCCTG 1593  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1594 GATTTCAGAAAGCAAGTCAGCTGTTTCAAGCTGCACCTGGAGAAGCGAGCAAGGTGCGACT 1653  
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1654 ACAATCGTGTAGCAGACCCGACTTCTTACTATTTCGAAGTGCAGATTGTATGTGACCTTA 1713  
Qy 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580

Db 1714 AAGATGGGAATGTCGGGAGAAAGAGACATGCTGAACATAATGGCAAAACAGAGTCTA 1773  
Qy 581 TTTTyrSerLeuValMetSerGlnAspIleLeuLysLeuAlaAspGluMetGluSerMet 600  
Db 1774 TATTATTCACTTGTGATGTCACAGGATATTAAGAAAGCTGATGAACAGATGGAGTCAATG 1833  
Qy 601 ThrTySerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620  
Db 1834 ACATATTCTACTGAAAGAAAGACCAACTCACTTCTCTGCACTCTGTGAAGACATCAAG 1893  
Qy 621 SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGlu 640  
Db 1894 TCAGACTTCAATGACAAAGGCTGAGGAATCCACCAATCTAAAGAGATAAGTCTTCCTGAA 1953  
Qy 641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660  
Db 1954 GTCTCTCTATAAANAATTTAAGATTAAACAAGCCTGAATGGCCTTTGTGGTCTGGGG 2013  
Qy 661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleLeuPheAlaLys 680  
Db 2014 ACATTGGCTTCTGTTCTAATGGAATGTTTCAATCCAGTATTTTCCATCATCTTTGCAAAA 2073  
Qy 681 IleIleThrMetPheGlyAsnAsnAspLysThrThreuleylshisaspAlaGluIleTyr 700  
Db 2074 ATTATTAACCATGTTTCGNAATATATGATAAACACACATTAAGCATGATGAGAAATTTAT 2133  
Qy 701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720  
Db 2134 TCCATGATATTCTGCTCAATTTGGGTGTTATTTGGCTTTGTCTAGTTATTTTCATGCAAGGATTA 2193  
Qy 721 PheTyrGlyVargAlaGlyIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740  
Db 2194 TTTTACGGCAGACAGCGGGAATTTTAAACCATGAGATTAAAGACACTTTGGCCTTCAAGCC 2253  
Qy 741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760  
Db 2254 ATGTTATATCAGGATATTGCTCGTGGTTGATGAAAGGAAACACAGCAGGAGGCTTGACA 2313  
Qy 761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780  
Db 2314 ACAATATTAGCCATAGATATAGCACAAATTAAGAGGCAACAGGTTCCAGGATGGCTC 2373  
Qy 781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 800  
Db 2374 TTAACACAAATGCAATCAATCAATGAGCTTTCACTTATCATCTTCTTATATATGATGATG 2433  
Qy 801 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820  
Db 2434 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTCCCGTGACAGGAATGATTGAA 2493  
Qy 821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840  
Db 2494 ACCGACGCAATGACTGGAATTTGCCAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553  
Qy 841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860  
Db 2554 ATAGCAACTGAACCTTTGGAGATATACGTACTATAGTGTTCATTAAACAGGGAAGGCC 2613  
Qy 861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880  
Db 2614 TTCGAGCAATGATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2673  
Qy 881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900  
Db 2674 CAGATATTGGAAGCTGTTATGATTCATTCAGCCATGCTCTTTATATATTGCTATGAGCA 2733  
Qy 901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920  
Db 2734 GGGTTTCGATTGGAGCTTATTAATTCAGCTGGAGGAATGACCCAGAGGCGCATGTTTC 2793  
Qy 921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940

Db 2794 ATAGTTTTTACTGCAATTCATATGAGAGCTATGGCATCGGAAAAACGCTGTTTTGGCT 2853  
Qy 941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960  
Db 2854 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTGGCTTTGTTGAAAAAGAAA 2913  
Qy 961 ProAsnIleAspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeu 980  
Db 2914 CCAAATATTAGACGCGCAGTCAAGAAAGGNAAGAACCCAGACACATGTGAAGGAATTTA 2973  
Qy 981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000  
Db 2974 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTGCGCCAGATGTTTTCATCTCCGTGGC 3033  
Qy 1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020  
Db 3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTTGTGGGAGCAGCGCTGTGGG 3093  
Qy 1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040  
Db 3094 AAAAGCACTTCTGTTCACCTTCTGCAGAGACTTATAGACCCCGTGCAGGACAAAGTGCTG 3153  
Qy 1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060  
Db 3154 TTTGATGGTGTGATGCCAAAGAAATTAATGATGATGAGTGCTCCGTTCCCAATATAGCAATC 3213  
Qy 1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080  
Db 3214 GTTCTCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTTGCTGAGAACATCGCCTATGTGAC 3273  
Qy 1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAsnAlaAlaIleHis 1100  
Db 3274 AACAGCGTGTGTGCCATTAGATGAGATCAAGAAAGCCGCAAAATGCAGCAATATCCAT 3333  
Qy 1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120  
Db 3334 TCTTTTATTGAAGGTCTCCCTGAGAAATACACACAAGTTGGACTGAAAGAGGACAG 3393  
Qy 1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140  
Db 3394 CTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453  
Qy 1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160  
Db 3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAGGTGGTTTCA 3513  
Qy 1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180  
Db 3514 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCTAGTGGTCACTCACAGGCTCTCT 3573  
Qy 1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200  
Db 3574 GCATTCAGAACGCGAGATTTGATGGTTCTGCACAATGGAAAGATAAAGGAAACAGGA 3633  
Qy 1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 3634 ACTCATCAAGAGCTCTCGAGAAATCGAGACATATATTTAAAGTTAGTGAATGCACAGTCA 3693  
Qy 1221 ValGln 1222  
Db 3694 GTGAC 3699

## RESULT 2

AX339032 LOCUS AX339032 3621 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 14 from Patent WO0194400.

ACCESSION AX339032

VERSION AX339032.1 GI:18129124

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



REFERENCE 1  
AUTHORS Frank, M.H. and Savegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 14 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES Location/Qualifiers  
source 1..3621  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1081 a 696 c 842 g 1002 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3621  
Score: 5988.50 Matches: 1195  
Percent Similarity: 97.79% Conservative: 0  
Best Local Similarity: 97.79% Mismatches: 0  
Query Match: 97.23% Indels: 27  
DB: 6 Gaps: 1  
US-09-873-409-5 (1-1222) x AX339032 (1-3621)  
QY 1 MetIleuLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 34 ATGATCTGGGTATACCTGGCATCACTGGTCAATGGAGCGCTTCCTTAAATGCCACTG 93  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 94 GTTTTAGGAGAAATAGTATGATTAATAGTGGATGTCTAGTCCAAACTAACACATAC 153  
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
DB 154 TCITTCCTCAGGTGACCCGTATATATGTGGAAATAGGTGTGCTGCCCTTGATTTTGGT 213  
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
DB 214 TACATACAGATTTCCTGTGGATTATTAATCGACACGACAGCAAGAGGATTCGAAAA 273  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
DB 274 CAGTTTTTTTCATTCAGTTTTGGCAGAGATCGGCTGGTTGATAGCTGTGACATCGGT 333  
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 334 GAACCTTAACCTCGCATGACAGACATTCACAAATCACTGATGGTATTCAGATTAAGATT 393  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 394 GCTCTGTTGTTTCAAAACATGCTCTACTTTTCGATTGGCCTGGCAGTTGTTGGTGAAG 453  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProIleuIleMetAlaSerAlaAla 160  
DB 454 GGCTGGAAACTCACCTAGTGACTCTATCCACGTCCTCTCTTATAATGGCTTCAGCGGCA 513  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 514 GCATGTTCTAGGATGGTCATCTCAATGACAGTAAGGAATTAAGTGCCTATTTCCAAAGCT 573  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 574 GGGCTGTGGCAGAGAGACTTGTCTCATCAATCCGACAGCTCATAGCTTTAGGGCCACG 633  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 634 GAGAAAGAACTTCAA----- 648  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 649 -----AGGTATACAGAAATCTCAAAGAT 672  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260

DB 673 GCAAGGATTTTGGCATAAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTAC 732  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
DB 733 TTTCTTTATGAATGGAACTTATGGACTTGTCTTTTGGTATGGAACTCTCTTGTATTTAAT 792  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 793 GGAGAACCTGGATATACCATCGGACTGTCTTGTCTGTCTTTCTTTAGTGTAAATCCATAGC 852  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 853 AGTTATTTGCAATGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCT 912  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 913 GCCTTTTCATATTTTCAGGTTTATGATAGAAACCCAGATATAGATACTTTCCACAGCT 972  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 973 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAATTTAAAAATGTTCTTCTTCAATTAT 1032  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1033 CCATCAAGACCATCTATCAAGATTTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAG 1092  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1093 ACAGTCGCTTGTGCTGCTCAATGGCAGTGGGAAAGATACGGTAGTCCAGCTTCTGCAG 1152  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
DB 1153 AGGTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTA 1212  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1213 AATGTGCGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1272  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 1273 ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1333 AGAGCAGCAGGGAACAAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1392  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1393 ACATTTGTTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAAT 1452  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 1453 GCTCTGTCCTTAGTTCGAAACCCCAAGATTTCTGATTTTATAGATGAGGCTAGCTCTGCCCTG 1512  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuLeuGluLysAlaSerLysGlyArgThr 540  
DB 1513 GATTCAAGAAACGAGTCAGCTGTTCAAGTGCCTGAGAGAGCGGAGCAAGGCTCGGACT 1572  
QY 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
DB 1573 ACAATCGTGTAGTAGCACCGACTTCTTACTATTCGAAGTGCAGATTTGATTTGTGACCCCTA 1632  
QY 561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580  
DB 1633 AAGGATGGAAATGCTGGCGGAGAAAGGAGCAGCATGCTGAACATATGCGCAAAACAGGCTCTA 1692  
QY 581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600  
DB 1693 TATATTTCACTTGTGATGTCACAGGATATTTAAAAAGCTGATCAACAGATGGAGTCAATG 1752  
QY 601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620  
DB 1753 ACATATTTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGGAAGATCATCAAG 1812



## FEATURES

## source

Location/Qualifiers

1. .3699  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/notes="Incyte ID No: 7472030CB1"

BASE COUNT 1116 a 707 c 860 g 1016 t

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 3699  
Score: 5737.00 Matches: 1156  
Percent Similarity: 93.78% Conservative: 4  
Best Local Similarity: 93.45% Mismatches: 9  
Query Match: 93.15% Indels: 68  
DB: 6 Gaps: 4

US-09-873-409-5 (1-1222) x AX478104 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 145 ATGATCTGGGTACTGACATCACTGTTCAATGGAGCCTTCCTTCTTAATGCCACTG 204  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGATGCTCTAGTCCACACTAACACA 264  
QY 41 SerPhePheArg-----LeuThr 46  
DB 265 AATTATCAGAACTGACTAGTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324  
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
DB 325 CTGTATTATGTGGAAATAGGTGTTGCTGCTTGTATTTTGGTTACATACAGATTTCTCTG 384  
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86  
DB 385 TGGATTATAACTGCAGCAGCAGACAGCAAGAGGATTCGAAACAGTTTTTTTCATTCAGTT 444  
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
DB 445 TTGSCACAGGACATCGGCTGGTTTGAAGCTGTGACATCGGTGAACCTTAACATCTCGCATG 504  
QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
DB 505 ACAGATCAGATTGCACAAATCAGTGATGGTATTGGACATAAGATTGCTCTGTTGTTTCAA 564  
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThr 145  
DB 565 AACATGCTACTTTTTCGATTGGCTGGCAGTTGGTTTGGTGAAGGCTGGAAACTCACC 624  
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165  
DB 625 CTAGTGACTCTATCCAGCTCTCTCTTATTAATGGCTTCAGCGGCAGCATGTTCTAGATG 684  
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
DB 685 GTCACTCTATGACCATTAAGGATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744  
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205  
DB 745 GAAGCTTGTTCATCAATCCGAAACAGTATAGCCTTTAGGCCCCAGGAGAAAGAACTTCAA 804  
QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225  
DB 804 ----- 804  
QY 226 SerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245  
DB 805 -----AGGTATACAGAAATCTCAAGATGCAAGGATTTTGGC 843  
QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265  
DB 844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTCTGGTGTGTGTACTTCTTTTATGAATGGA 903

QY 266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285  
DB 904 ACCTATGGACTTGTCTTTTGGTATGGAACTCTCTGATTCTTAATGGAGAACTGGATAT 963  
QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305  
DB 964 ACCATCGGACTGTTCTTCTGCTGTTTCTTTAGTGTAAATCCATAGTAGTATTATTGCA 1023  
QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325  
DB 1024 GCAGCAGTCCCTCCTTGTAAACCTTCGCAATAGCCGAGGAGCTGCCCTTCATATTTTC 1083  
QY 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345  
DB 1084 CAGGTTATTGATPAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAACCTGAA 1143  
QY 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365  
DB 1144 TCCATAGAAGGAACCTGGAATTTAAAAATGTTTCTTTCAATTTATCCATCAAGGCCATCT 1203  
QY 366 IleLysIleLeuLysGlyLeuAsnLeuArgLysSerGlyGluThrValAlaLeuVal 385  
DB 1204 ATCAAGATTCTGAAGAGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCTTGGTC 1263  
QY 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405  
DB 1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCGACAGAGTTATATGATCCG 1323  
QY 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425  
DB 1324 GATGATGGCTTTATCATCTGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCAT 1383  
QY 426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445  
DB 1384 CGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAC 1443  
QY 446 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465  
DB 1444 AATATCAAGTATGGACAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGAA 1503  
QY 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485  
DB 1504 GCAAAATCGCTATGATTTTATCATGGAGTTCCTTAATAAAATTTAATCAATTTGGTAGG 1563  
QY 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505  
DB 1564 AAAGGAGCTCAATGATGAGTGGAGGCGAGAAACAGAGATTCGCAATTCGCTGCTTAGTT 1623  
QY 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525  
DB 1624 CGAAACCCCAAGATTTCTGATTTTAGATGAGGCTACGCTGCCCCCTGGATTCCAGAA 1683  
QY 526 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545  
DB 1684 TCAGCTGTTCAAGCTGCACCTGGAGAAAGGCGAGCAAAAGGTCGAGCTCAATTCGTGTAG 1743  
QY 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565  
DB 1744 CACCGACTTCTACTATTGCAAGTGCAGATTTGATTTGTGACCCCTAAAGGATGAATGCTG 1803  
QY 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585  
DB 1804 CGCGAGAAAGGAGCGCATGCTGAACCTAATGGCAAAACGAGGTCTATATTATTCACTTGT 1863  
QY 586 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 605  
DB 1864 ATGTCACAGGATATTAATAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAA 1923  
QY 606 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625  
DB 1924 AGAAAGCAACCACTCCTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGAC 1983

```
QY 626 LyeAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLys 645
DB 1984 AAGCGTGAGGAATCCACCAATCTAAAGAGATAAGTCTTCCCTGAAGTCTCTCTATTAAAA 2043
QY 646 IleLeuLysLeuAsnLysProGluTppProPheValValLeuGlyThrLeuAlaSerVal 665
DB 2044 ATTTTAAAGTTAAACAAGCTGAATCGCCTTTTGTTGTTCTGGGACATTTGGCTTCTGTT 2103
QY 666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleThrMetPhe 685
DB 2104 CTAAATGGAACTGTTTCATCCAGTATTTCATCATCTTTGCAAAAAATTATAACCATGTTT 2163
QY 686 GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
DB 2164 GGAATAATGATAAAACCAATTAAGAGCATGATGCAAGAAATTTATTCCATGATATTTCGTC 2223
QY 706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
DB 2224 ATTTTGGGTGTTATTGCTTTTGCTTTCAGTATTTCATG----- 2259
QY 726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
DB 2260 -----CAGGAT 2265
QY 746 IleAlaTppPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
DB 2266 ATTGCTGCTGTTTGTATGAAAGGAAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATA 2325
QY 766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
DB 2326 GATATAGCAAAATTCAGAGGCACAGGTTCCAGGATTGGCGTCTTTAAACAAAAATGCA 2385
QY 786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTppGluMetThrPheLeu 805
DB 2386 ACTAACATGGGACTTTTCAGTTATTCATTCTCTTATATATGATGGAGATGACATTTCCTG 2445
QY 806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaIleMetThr 825
DB 2446 ATTTCTGAGTATTGCTCCAGTACTTGCCTGACAGGAATGATTGAAACCCGACGCAATGACT 2505
QY 826 GlyPheAlaAsnLysAspLysGlnGluLysHisAlaGlyLysIleAlaThrGluAla 845
DB 2506 GGATTTGCCAACAAAGATAAAGCAAGCACTTAAGCATGCTGGAAAGATAGCAACTGAAGCT 2565
QY 846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
DB 2566 TTGGAGAATATACGTACTATAGTGTCATTAAACAGGAAAGCCCTTCGAGCAATGTAT 2625
QY 866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysAlaGlnIleIleGlySer 885
DB 2626 GAAGAGATGCTTCAGACTCAACACAGAAAATACCTCGAAGAAAGCACAGATTATTGGAAGC 2685
QY 886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
DB 2686 TGTATTGCAATTCAGCCATGCCCTTTATATATTGTCCTATGCGCAGGGTTTCGATTGGA 2745
QY 906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
DB 2746 GCCTATTTAATTCAGCTCGAGCAATGACCCAGAGGGCATGTTCTAGTATTTTACTGCA 2805
QY 926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
DB 2806 ATTGCAATATGGAGCTATGGCATCGGAAACCGCTCGTTTGGCTCCTCAATATTCCAAA 2865
QY 946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
DB 2866 GCCAAATCGGGGGCTGCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGC 2925
QY 966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
DB 2926 CCGAGTCAAGAGGGAAGAACGACACATGTGAAGGAATTTAGAGTTTCAGAAAGTC 2985
QY 986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
```

```
DB 2986 TCTTTCTCTATCCATGTCGCCAGATGTTTTTCATCCTCCGGGCTTATCCCTCAGTATT 3045
QY 1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
DB 3046 GAGCGAGGAAAGACAGTACATTTTGTGGGAGCAGCGGCTGTGGGAAAGACATTTCTGTT 3105
QY 1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyValLeuPheAspGlyValAsp 1045
DB 3106 CCACTTCTCAGAGACTTTATGACCCCGTCAAGGACAAAGTGTGTTTGTATGTTGGAT 3165
QY 1046 AlaLysGluLeuAsnValGlnTppLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
DB 3166 GCAAAAGAAATTGAATGTACAGTGGCTCCGTTCCTCAATAGCAATTCCTCAAGAGCCT 3225
QY 1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
DB 3226 GTGCTCTTCAACTGCAGCATTTGCTGGAACATCGCTATGTTGACAAACAGCCGTGGTG 3285
QY 1086 ProLeuAspGluIleLysGluAlaAlaAsnIleHisSerPheIleGluGly 1105
DB 3286 CCATTAGATGAGATCAAGAGCCGCAATGCGCAAAATATCCATTCTTTTATTGAAGGT 3345
QY 1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
DB 3346 CTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGAGCACAGCTTTCTGCGGCCAG 3405
QY 1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAsp 1145
DB 3406 AAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTGGAT 3465
QY 1146 GluAlaThrSerAlaLeuAspAsnAspSerGluValValGlnHisAlaLeuAspLys 1165
DB 3466 GAGGCCACTTTCAGCCCTCGATAATGACAGTGAAGAGTGTTCAGCATGCCCTTGATAAA 3525
QY 1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
DB 3526 GCCAGACGGAAGGACATGCTTAGTGTCCTCAGCTCTCTGCAATTCAGAACGCA 3585
QY 1186 AspLeuIleValValLeuHisAsnGlyLysIleLysValGlnGlyThrHisGlnGluLeu 1205
DB 3586 GATTGTAGTAGTGGTCTGCACAATGGAAGATAAAGAAACAAGGAACCTCATCAAGAGCTC 3645
QY 1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
DB 3646 CTGAGAAATCGAGACATATATTTTAAATTAGTTAGTGAATGACAGTCAGTCAG 3696
RESULT 4
AX339030 3177 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 12 from Patent WO0194400.
DEFINITION AX339030
ACCESSION AX339030
VERSION AX339030.1 GI:18129122
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Frank,M.H. and Sayegh,M.H.
TITLE A gene encoding a multidrug resistance human p-glycoprotein
homologue on chromosome 7p15-21 and uses thereof
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES
source
1. .3177
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 970 a 601 c 735 g 870 t 1 others
ORIGIN
Alignment Scores: 0 Length: 3177
Pred. No.:
```

Score: 5328.00 Matches: 1058  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 86.51% Indels: 0  
DB: 6 Gaps: 0

US-09-873-409-5 (1-1222) x AX339030 (1-3177)

QY 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184  
DB 1 ATGGTCATCTCAATGACAGTAAGAAATTAAGTCCCTATTCCAAAGCTGGGCTGGCA 60  
QY 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204  
DB 61 GAAGAAGCTTGTTCATCAATCCGAACAGTCATAGCCCTTAGGGCCCCAGAGAAAGAACTT 120  
QY 205 GlnArgSerPheLeuLeuAenIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224  
DB 121 CAAGAAGCTTCTCTTTAAATATATAACAAGATATGCTTTGGTTTATTTTCCCCAGTGGCTA 180  
QY 225 LeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAenLeuLysAspAlaLysAspPhe 244  
DB 181 CTAAGTGTGTCTGTNTTGTAAAGGTATACACAGAAATCTCAAGATGCCAAGGATTTT 240  
QY 245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264  
DB 241 GGCATAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTACTTCTTTATGAAT 300  
QY 265 GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAenGlyGluProGly 284  
DB 301 GGAACCTATGACATGCTTTTGGTATGGAACCTCTTGTGATTTCTTAATGGAGAACCTGGA 360  
QY 285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304  
DB 361 TATACCATCGGAGCTGTCTTGTGCTTTCTTTAGTGTAAATCCATAGCAGTATTGCAAT 420  
QY 305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324  
DB 421 GGACGAGCAGTCCCTCACTTTGAAACCTTCCCAATAGCCCGAGGAGCTGCTTTCATATT 480  
QY 325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344  
DB 481 TTCAGAGTATTGTATAAGAACCCAGCATAGATAACTTTTCCACAGCTGGATATAAACCT 540  
QY 345 GluSerIleGluGlyThrValGluPheLysAenValSerPheAenTyrProSerArgPro 364  
DB 541 GAATCCATAGAGGAACGTGTGGAATTTAAAAATGTTCTTTCATATATCCATCAAGACCA 600  
QY 365 SerIleLysIleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeu 384  
DB 601 TCTATCAAGATTCTCAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCTTG 660  
QY 385 ValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404  
DB 661 GTCGCTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATGAT 720  
QY 405 ProAspAspGlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHis 424  
DB 721 CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGGAT 780  
QY 425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444  
DB 781 TATCGAGACCATATTTGGAGTGGTAGTCAAGAGCTGTTTGTTCGGGACCCACCATCAGT 840  
QY 445 AenAenIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464  
DB 841 AACCAATATCAAGTATGGACGAGATGATGTGATGATGAAGAGATGGAGAGAGCGACGAGG 900  
QY 465 GluAlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGly 484  
DB 901 GAAGCAAAATGCGTATGATTTATATCATGAGGTTTCTTAATAAATTTAATACATTTGTGAGG 960  
QY 485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504

DB 961 GAAAAAGGAGCTCAAAATGATGGAGGCGAAAAACAGAGGATCGCAATGCTCTGTGCTTTA 1020  
QY 505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524  
DB 1021 GTTCGAACCCCAAGATTTGATTTTATAGTGGCTACGCTCTGCCCTGGATTTCAGAAAGC 1080  
QY 525 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 544  
DB 1081 AAGTCAGCTGTTCAAGCTGCACCTGGAGAGGCGAGCAAGGTCGACATCAAAATCGTGGTA 1140  
QY 545 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 564  
DB 1141 GCACACGACTTTCTATATTCCGAAGTGCAGATTTGATGTGACCTTAAAGGATGGAAATG 1200  
QY 565 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 584  
DB 1201 CTGGCGAGAAAGAGCACATGCTGAACCTAATGGCAAAACAGAGTCTATATTATTCATT 1260  
QY 585 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 604  
DB 1261 GTGATGTCCAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320  
QY 605 GluArgLysThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624  
DB 1321 GAAAGAAAGACCAACTCACTTCTCTGCACTGTGTGAAGAGCATCAAGTCAGACTTCTATT 1380  
QY 625 AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644  
DB 1381 GACAAGCTCAGGAATCCCAACTCTAAAGAGATAGTCTTCTGAAAGTCTCTCTATTATA 1440  
QY 645 LysIleLeuLysLeuAenLysProGluTrpPheValValLeuGlyThrLeuAlaSer 664  
DB 1441 AAAATTTAAAGTTAAACAAGCTGAATGCCCTTTGTGTTCTGGGGACATGGCTTCT 1500  
QY 665 ValLeuAenGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684  
DB 1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCCAAATAATATACCATG 1560  
QY 685 PheGlyAenAenAspLysThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704  
DB 1561 TTTGGAAATTAATGATAAAACCAATTAAGCATGATGACAGAAATTTATTTCCATGATATTC 1620  
QY 705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724  
DB 1621 GTCATTTTGGTGTATTGCTTGTTCAGTTATTTTCATGAGGAGTATTTTACGGCAGA 1680  
QY 725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744  
DB 1681 GCAGGGGAAATTTTAAAGATGAGATTAAAGACACTTGGCTTCAAGGCTATTTATATCAG 1740  
QY 745 AspIleAlaTrpPheAspGluLysGluAenSerThrGlyGlyLeuThrThrIleLeuAla 764  
DB 1741 GATATTGCTGTTTGTATGAAGAAAGAAACAGCACAGAGGCTTGACACAAATATTAGCC 1800  
QY 765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAen 784  
DB 1801 ATAGATATAGCAAAATTTCAAGGAGCAACAGGTTCCAGGATTGGCTCTTAAACAAAAA 1860  
QY 785 AlaThrAenMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804  
DB 1861 GCAACTAACATGGGACTTTCAGTTTATCATTTCTTTATATATGATGGAGATGACATTC 1920  
QY 805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824  
DB 1921 CTGATTTCTGAGTATTTGCTCCAGTACTTGGCTGACAGAAATGATTAAGACCGCAATG 1980  
QY 825 ThrGlyPheAlaAenLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844  
DB 1981 ACTGGAATTCGCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAGATAGCAACTGAA 2040  
QY 845 AlaLeuGluAenIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864

Db 2041 GCTTTGGAGATATACGTACTATAGTGTCTATTAAACAGGAGAAAGCCCTTCGAGCAATG 2100

Qy 865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884

Db 2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAGACACAGATATTGGA 2160

Qy 885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904

Db 2161 AGCTGTATTGCATTACGCTTATATATATTTTGGCTTATGCACAGGGTTTCGATTT 2220

Qy 905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924

Db 2221 GGAGCCTATTATTCAGCTGAGCGAATGACCCAGAGGGCAGTTCATAGTTTTTACT 2280

Qy 925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944

Db 2281 GCAATTGCATATGAGCTATGGCCATCGGAAACGCTCGTTTTGGCTCCTGAATATTC 2340

Qy 945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964

Db 2341 AAAGCCAAATCGGGGCTGCGCATCTGTTGGCTTGTGGAAAGAAACCAATATAGAC 2400

Qy 965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984

Db 2401 AGCCGAGCTCAGAGAGGAAAGAACCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAA 2460

Qy 985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004

Db 2461 GTCTCTTTCTTATCATGTCGCCCAGATGTTTTTCATCTCGTGGCTTATCCCTCAGT 2520

Qy 1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024

Db 2521 ATTAGCGAGGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGCACTTCT 2580

Qy 1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1044

Db 2581 GTTCAACTTCGAGAGACTTATGACCCCGTGCAAGGCAAGTGCTGTTGATGTGTG 2640

Qy 1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064

Db 2641 GATGCAAGAGATTGAATGTACAGTGGCTCCGTTCCCAATAGCANCTGTTCTCAGAG 2700

Qy 1065 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1084

Db 2701 CCGTGTCTCTTCAACTGCAGCATTGTGAGAACATCGCCTATGTTGACAAACAGCCGTGT 2760

Qy 1085 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGlu 1104

Db 2761 GTGCCATTAGATGAGATCAAGAAAGCCGCAATGCGAGCAATATCCATTTTATTGAA 2820

Qy 1105 GlyLeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGly 1124

Db 2821 GGTCTCCCTGAGAAATACACACACAGTTGGACTGAAGGAGCACAGCTTTCGGCGGC 2880

Qy 1125 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1144

Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGCTCTTCTCCAAAAACCCCAAAATTTATTGTTG 2940

Qy 1145 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1164

Db 2941 GATGAGGCCACTTCAGCCCTCGAATAATGACGTGAAGGTGGTTTCAGATGCCCTTGTAT 3000

Qy 1165 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1184

Db 3001 AAAGCCAGAGGGGAGGACATGCTAGTGGTCACCTCAGAGGCTCTCTGCAATTCAGAAC 3060

Qy 1185 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1204

Db 3061 GCAGATTTCATAGTGTCTTGCACAATGGAAGATAAAGGAAACCAAGAACTCATCAAGAG 3120

Qy 1205 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222

Db 3121 CTCCTGAGAAATCGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTCGAG 3174

RESULT 5  
AX339028  
LOCUS AX339028 2856 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 10 from Patent WO0194400.  
ACCESSION AX339028  
VERSION AX339028.1 GI:18129120  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Frank,M.H. and Sayegh,M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 10 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source location/Qualifiers  
1..2856  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 888 a 540 c 652 g 776 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.27e-296 Length: 2856  
Score: 4561.50 Matches: 912  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 74.06% Indels: 18  
DB: 6 Gaps: 1  
US-09-873-409-5 (1-1222) x AX339028 (1-2856)  
Qy 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheG1 312  
Db 114 GTTTCCTTTAGTGTAAATCCATAGCAGTTATTTCATTGGAGCAGCAGTCCCTCATTTATT-- 171  
Qy 312 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 332  
Db 172 -----GATAAGAAAC 182  
Qy 332 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValG1 352  
Db 183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGGAACCTGTGA 242  
Qy 352 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 372  
Db 243 ATTTAAAAATGTTCTTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302  
Qy 372 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 392  
Db 303 GAATCTCAGAAATTAAGTCTGGAGACAGCTCCCTTGGTGGTCTCAATGGCAGTGGAA 362  
Qy 392 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 412  
Db 363 GAGTACGGTAGTCCAGCTTCTCAGAGGGTTATATATCGGATGATGGCTTTATCATGT 422  
Qy 412 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 432  
Db 423 GGATGAGAATGCATCAGAGCTTTAAATGTGGCGCATTTATCGAGACCATATTTGGAGTGT 482  
Qy 432 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 452  
Db 483 TAGTCAAGAGCTGTTTGTTCGGACCCACCATCAGTAAACATATCAAGTATGGAGAGA 542  
Qy 452 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheI1 472  
Db 543 TGATGTGACTGATGAAGATGAGAGAGCAGCAGCAAGGAGCAAAATCGTATGATTTTAT 602  
Qy 472 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerG1 492



Db 603 CATGGAGTTCCTAATAATTTAATCATTTGGTAGGGGAAAAAGAGCTCAAAATGAGTGG 662  
Qy 492 YGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuI 512  
Db 663 AGGCAGAAACAGAGATCGCAATTGCTCGCTTAGTTCGAAACCCCAAGATTCTGAT 722  
Qy 512 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLe 532  
Db 723 TTTAGATGAGGCTACGCTCGCTCGCTCGATTGAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782  
Qy 532 uGluLysAlaSerLysGlyArgThrThrIleValValAlaAlaHisArgLeuSerThrIleAr 552  
Db 783 GGAAGAGCCGAGCAAGGTCGAGTACAAATCGTGTAGCACACCGACTTTCTACTATTTC 842  
Qy 552 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 572  
Db 843 AAGTGCAGATTGATTGTGACCCCTAAAGGATGGATGCTGGCGAGAAAGAGACACATGC 902  
Qy 572 aGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSerGlnAspIleLysLy 592  
Db 903 TGAACATAATGGCAAAACAGGCTCTATTATTTCATTGTGATGTCACAGGATATTAATAA 962  
Qy 592 sAlaAspGluGlnMetGlnSerMetThrTySerThrGluArgLysThrAsnSerLeuPr 612  
Db 963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022  
Qy 612 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluSerThrGl 632  
Db 1023 TCTGCATCTGTGAAAGACATCAAGTCAGACTTCATTGACACAGGCTGAGGAATCCACCCA 1082  
Qy 632 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 652  
Db 1083 ATCTAAAGAGATAGTCTTCTGTAAGTCTCTCTATTATAAATAATTTAAAGTTAAACAAGCC 1142  
Qy 652 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 672  
Db 1143 TGAATGGCCCTTTGTGTCTCGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202  
Qy 672 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 692  
Db 1203 AGTATTTTCCATCATCTTTGCATAAATAATTAACCATGTTTGGAAATTAATGATTAACACAC 1262  
Qy 692 rLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPh 712  
Db 1263 ATTAAGCATGATGCAGAAATTTATTCCATGATATTGTCATTTGGTGTATTGCTT 1322  
Qy 712 eValSerTyTrpPheMetGlnGlyLeuPheTyArgAlaGlyGluIleLeuThrMetAr 732  
Db 1323 TGTCAGTTATTTCATGCGAGGATTTATTTACGGCAGACGAGGGGAAATTTTAACGATGAG 1382  
Qy 732 gLeuArgHisLeuAlaPheLysAlaMetLeuTyTrpGlnAspIleAlaTrpPheAspGluLy 752  
Db 1383 ATTAAGACACTTGGCCCTTCAAGCCATGTTATATACAGATATTCCTGGTTTGTGATGAAA 1442  
Qy 752 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 772  
Db 1443 GGAATAACAGCACAGAGGCTTGACAAATATTAGCCATAGATATAGCAACAAATTCAGG 1502  
Qy 772 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 792  
Db 1503 AGCAACAGGTTCCAGGATGCGCTTTAAACACAAATGCAACTAACATGGGACTTTTCAGT 1562  
Qy 792 lIleIleSerPheIleTyTrpGlyMetThrPheLeuIleLeuSerIleAlaProVa 812  
Db 1563 TATCATTTCTTTATATATGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622  
Qy 812 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 832  
Db 1623 ACTTGGCTGACAGGAATGATTGAACCGCAGCAATGACTGATTTGCCAACAAAGATAA 1682  
Qy 832 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 852  
Db 1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTTGAGAAATATACGTACTAT 1742

Qy 852 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyTrpGluGluMetLeuGlnThrGl 872  
Db 1743 AGTGTCTATTAAACAAGGCAAAAGCCCTTCGAGCAAAATGTTATGAAGAGATGCTTCAGACTCA 1802  
Qy 872 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyTrpAlaPheSerHisAl 892  
Db 1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTATTGCAATTCAGCATTGCG 1862  
Qy 892 aPheIleTyTrpPheAlaTyTrpAlaAlaGlyPheArgPheGlyAlaTyTrpLeuIleGlnAlaGl 912  
Db 1863 CTTTATATATTTTGCCTATGACAGAGGTTTCGATTGGAGGCTATTATTAATTCAGACTGG 1922  
Qy 912 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyTrpGlyAlaMetAl 932  
Db 1923 ACGAATGACCCACAGAGGCGATGTTTCATAGTCTTTTACTGCAATTCGATATGGAGCTATGGC 1982  
Qy 932 alleGlyLysThrLeuValLeuAlaProGluTyTrpSerLysAlaLysSerGlyAlaAlaHi 952  
Db 1983 CATCGGAAAAACGCTCGTTTTTGGCTCTCTGAATATTCCAAAGCCAAATCGGGGCTCGCA 2042  
Qy 952 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 972  
Db 2043 TCTGTTTGGCTTGTGGAAGAAAGAAACCAATATAGACAGCCGCGTCAAGAGGGAAGAAA 2102  
Qy 972 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyTrpProCysAr 992  
Db 2103 GCCCAGACACATGGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTCTATCATGTCG 2162  
Qy 992 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 1012  
Db 2163 CCCAGATGTTTTTCATCTCCGCTGCTTATCCCTCAGTATTGAGCGAGGAAAGACATGAGC 2222  
Qy 1012 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1032  
Db 2223 ATTTTGGGGAGGAGCGGCTGTGGGAAAGACATCTCTGTTCAACTTCTGCAGAGACTTTA 2282  
Qy 1032 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1052  
Db 2283 TGACCCCGTGCAAGGACAAAGTGTCTGTTGATGTTGTTGATGCAAAAGAAATTAAGATGTACA 2342  
Qy 1052 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1072  
Db 2343 GTGGCTCCGTTCCAAATAGCAATCGTTCCTCAAGAGCCTGTCTCTTCAACTGCGAGCAT 2402  
Qy 1072 eAlaGluAsnIleAlaTyTrpGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1092  
Db 2403 TGTGAGAACATCGCTTATGTTGACCAACAGCCGTGTGTGCTCCATTAGATGAGATCAAAGA 2462  
Qy 1092 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyTrpAsnTh 1112  
Db 2463 AGCCGCAATGCAAGCAATATATCCATCTTTTATTGAAGGTCTCCCTGAGAAAATACAACAC 2522  
Qy 1112 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1132  
Db 2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCGCAGAAAACAAAGACTAGCTATTGC 2582  
Qy 1132 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 1152  
Db 2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGATGAGGCCACTTCAGCCCTCGA 2642  
Qy 1152 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1172  
Db 2643 TAATGACAGTGAGAGGTGGTTTCAGCATGCCCTTGATAAAGCCAGGACGGAAGACATG 2702  
Qy 1172 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1192  
Db 2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCGCAGATTGTGATGTGTTCTGCA 2762  
Qy 1192 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1212  
Db 2763 CAATGGAAAGATTAAGGAACACAGGAACCTCATCAAGAGCTCTCGAGAAATCGAGACATATA 2822



[illegible]

Db	250	ATGCTGGTGGGAGACTCTCCGACGATATATATCCATGAGAGTTGCATCTCCCGCTTATGATGCTG	309
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn	38
Db	310	GTGTTTGGAGACATGACACACAGCTTTGCAAGTGTAGGAACAATCCCAACAATGCTACT	369
Qy	38	-----	38
Db	370	AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC	429
Qy	39	ThrTyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIle	58
Db	430	ACGTAGCCCTAT-----TACTACACAGGAGTTGGTGGTGGTGTCTTATA	474
Qy	59	PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle	78
Db	475	GTGTCCTACATTCAGGTTTCATTTTGGTCCTGGCAGCTGGGAACAATACACAAAT	534
Qy	79	ArgGlyGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp	98
Db	535	AGACAGAAGTTTTTTCATGCTATAATGAATCAGGAGATAGGTGCTTTCAGCTGCACGAC	594
Qy	99	IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly	117
Db	595	GTGCGGAGCTCAACACCCGGCTCAGATGATGCTCCAAATTAACGAAAGGAATTCGT	654
Qy	118	AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly	137
Db	655	GACAAATTTGGAATGTTCTTTCAGGCNATGGCAACATTTTGGTGGTTTTATAATAGA	714
Qy	138	LeuValIleGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla	157
Db	715	TTTACTCGAAGGCTGGAAGCTAAACCTTGTCATTTTGGCCATCAGCCCTGTTCTTGGACTG	774
Qy	158	SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr	177
Db	775	TCAGCTGGTATTTGGGCAAGATATATCTTCATTTACTGATAAAGAAGCTTCAGGCATAT	834
Qy	178	SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe	197
Db	835	GCAAAAGCTGGAGCAGTTGCTCAGAGAAGTCTTAGCAGCCATCAGAACTGTGATTCATTT	894
Qy	198	ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp	217
Db	895	GGAGGACAAAAGAAAGAACTTGAA-----	918
Qy	218	PheTyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsn	237
Db	919	-----AGGTACAACAACAAT	933
Qy	238	LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly	257
Db	934	TTAGAAGAAGCCAAAGAAGCTTGGGATAAAGAAGCGATCAGCCCAACATTTCCATGGGC	993
Qy	258	AlaValTyrPhePheMetAsnGlyTyrTrpGlyLeuAlaPheTrpTyrGlyThrSerLeu	277
Db	994	GCAGCTTCTCTTATCTATCATCATATGCTCTGGCCTCTCTGTAATGGAAGCTCTTG	1053
Qy	278	IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal	297
Db	1054	GTCACTCTCAAAAGAA-----TATTCTATTTGACAAGTGCCTCAGCTGCTCTCTTTGCTGTA	1107
Qy	298	IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla	317
Db	1108	TTAATTTGGGCGCATTCAGTATTTGGACAGGCATCTCCAAATATCGAAGCCCTTTGCCAATGCA	1167
Qy	318	ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe	337
Db	1168	AGAGGAGCAGCCTATGAATCTTCAACATAATTGATAATAGCCCCAGTATTGACAGCTTC	1227
Qy	338	SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer	357
Db	1228	TCAAGAATGGGTACAAACACAGACAACATTAAAGGAATTTGGAATTCAAAATATTCAAC	1287

QY 358 PheAsnTyrProSerArgProSerIleLeuLysGlyLeuAsnLeuArgIleLys 377  
 Db 1288 TTCACTTACCACTCGAAGAGAGCTTCAGATCTTGAGGCGCTCAACCTGAGAGTGCAG 1347  
 QY 378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397  
 Db 1348 AGCGACAGACAGTGGCCCTGGTGGCAACAGTGGCTGTGGGAAAGACCACTGTCCAG 1407  
 QY 398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417  
 Db 1408 CTCGTGACAGAGCTTACGACCCACAGAGCGGTGTGCTAGTATCGACGGTCAGGACATC 1467  
 QY 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437  
 Db 1468 AGGACCATCAATGTGAGTATCTCGGGAATCATTTGGGTGTGAGTCAGGAACCTGTG 1527  
 QY 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457  
 Db 1528 TTGTTTCCCACTATCGCTGAAACCATTCGCTATGCGCGAGAAAATGTCACTATGGAC 1587  
 QY 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477  
 Db 1588 GAGATCGAGAAAGCTGTCAGAGGAAGCAATGCCTATGACTTCATGAAACTGCCCCAT 1647  
 QY 478 LysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArg 497  
 Db 1648 AATTTGACACTCTTGTGTGTGAGAGAGGACACAGCTGAGTGGAGACAGAAACAGAGA 1707  
 QY 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuLeuAspGluAlaThr 517  
 Db 1708 ATGCCATTGCTGTCGCCCTGTCGCGCAACCCCAAGATCCTTTTGTGGATGAGCGACA 1767  
 QY 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLys 537  
 Db 1768 TCAGCCTTGGACACAGAAAGTGAAGCGGTGTCAGGCTGCTCGGATAGGCTAGAGAA 1827  
 QY 538 GlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557  
 Db 1828 GCGCGACTACCATGTGTATAGCCACCGCTGTCTACAGTTCGAAATGTGACATCAT 1887  
 QY 558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577  
 Db 1888 GCTGGGTTTGATGGTGTGTCATTGTGGAGCAAGAAATCATGAAGAGCTCATGAGAGAG 1947  
 QY 578 ArgGlyLeuTyrTyrSerLeuValMetSerGln----- 588  
 Db 1948 AAGGGCATTTTACTTCAAACTTGTGCATGACACAGACAGCAGCAAGAAATGAAATTGAAATTAGGA 2007  
 QY 589 ---AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 607  
 Db 2008 AATGAAGTGGTGTAGCTTAAAAATGAAATGAAATGAAATTTAGACATGCTTCAAAAGATTCA 2067  
 QY 608 ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625  
 Db 2068 GCATCCAGTCTAATCAGAAAGAGATCAACTCGCAGAAGTATCCGTGACACACATGACCAA 2127  
 QY 626 LysAlaGluGluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeu 644  
 Db 2128 GACAGGAAGCTTAGTACCAAGAGCGCTTGGATGAAGATGTACCTCCAAATTTCTTTTGG 2187  
 QY 645 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 664  
 Db 2188 CGTATCTTAAAGTTCAATTCATCTGAATGGCCATATTTTGGTGTGTATATCTGTGGC 2247  
 QY 665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684  
 Db 2248 ATAGTAAATGGAGCTTGAACACAGCATCTCTCAATAATATATTTCTAAGTTGTAGGGGT 2307  
 QY 685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703  
 Db 2308 TTCACAAAGAAATACTGTATGATGAACCAACCAACGACATGATGACAACTTGTTCCTTATTG 2367

QY 704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723  
 Db 2368 TTCTCATCTTGGGTCTATTTCTTTATTAATCTTTTCTTCTCAGGGCTTCACATTTGGC 2427  
 QY 724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743  
 Db 2428 AAGCTTGGAGAGATCTCCCAAGCGACTCCGATACATGGTTTCAAAATCCATGCTTGAGA 2487  
 QY 744 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763  
 Db 2488 CAGATGTGAGCTGGTGTGATTAACCCCTAAACACACCTGGACATGACCACAGGCTC 2547  
 QY 764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783  
 Db 2548 GCCAACGATCTGGTCAAGTTAAAGGGGCTACAGGAGCGAGACTTGTGTCAATACCCAG 2607  
 QY 784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 803  
 Db 2608 AATATAGCAAAATCTTGGGACAGGAATCATATCCTAATCTATGGCTGGCAGTTGACA 2667  
 QY 804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823  
 Db 2668 CTTCTACTCTTGAATTTGTCCTCCATCATTTGCNAATAGCAGGAGTGGTGGATGAAAAATG 2727  
 QY 824 MetThrGlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThr 843  
 Db 2728 TTGCTTGGCAAGCACTAAAAGATAGAGAGGAGCTAGAAAGTTCTGGGAAGATTGCTACT 2787  
 QY 844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863  
 Db 2788 GAAGCAATAGAGAACTTCCGCACCTGTCTCTTCTTACCTCGGAGAGCAAGTTTGAANAAT 2847  
 QY 864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883  
 Db 2848 ATGTATGCCAGAGCTTGCAGATACCATACAGAAATGCTCTGAAAGAGACACAGCTTTT 2907  
 QY 884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903  
 Db 2908 GGGATCACATTCCTTCCACCGCCATGATGATATTTTTCGTATGCTGTCTTGTTCGCG 2967  
 QY 904 PheGlyAlaTyrIleLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923  
 Db 2968 TTTGGTGTCTTACTTGTGGCAGCGGAATTAATGACATTTTGAANAATGTTCTATTAGTATTC 3027  
 QY 924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943  
 Db 3028 TCAGCTATTGCTTTTGTGCCATGGCAGTGGCAGTGGCAGTTCATTTCTCTCTGACTAT 3087  
 QY 944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsnIle 963  
 Db 3088 GCCAAAGCCAAAGTGTGGCATCGCATATCATCATGATCATGAAAAAGTCCCTTCCATT 3147  
 QY 964 AspSerArgSerGlnGluLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983  
 Db 3148 GACAGCTACAGCACCGGAGCGCTTAAGCCCTTAATACATTTGGAAGGAAATGGAANAATTAAT 3207  
 QY 984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003  
 Db 3208 GAAGTGTGTTTCAACTACCCCGCCCTGACATCCAGTGTCTTCAAGGCGCTGAACCTG 3267  
 QY 1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023  
 Db 3268 GAGTGAAGAAGGCGCAGACGCTGGCCCTTGTGGCAGCAGTGGCTGCGGGAAGAGACACA 3327  
 QY 1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043  
 Db 3328 GTGGTCCAGCTGTCTGAGCGGTTCTACAGCCCATCCCATCGCGGAACAGTGTTCATAGTGGC 3387  
 QY 1044 ValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGln 1063  
 Db 3388 AAGAGTAAACCACTAAATGTCCAGTGGCTCCGAGCACACCTAGGCATTTGTGCCAA 3447  
 QY 1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083



Db 903 GGAGGCAAAAGAAAGAACTTGAA----- 926  
Qy 218 PheTyrPheProGlnTyrPheLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsn 237  
Db 927 -----AGGTACAACACNAT 941  
Qy 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257  
Db 942 TTGAAGCAAGCCAAAGACTTGGGATAAAGAAAGCGATCACAGCCCAATTCATCGATGGC 1001  
Qy 258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeu 277  
Db 1002 GCAGCTTCTCTTATATATGATCATATCTCTGCGCTTCTGGTATGAAGCGTCTTG 1061  
Qy 278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297  
Db 1062 GTACATCTCAAAAGNA-----TAATCTATTGGACAAGTGTCTACTGTCTTCTTGGTGA 1115  
Qy 298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317  
Db 1116 TTAATTGCGCCATTCAGTATTGGACAGGCATCTCCAAATATCGAAGCTTTGGCAATGCA 1175  
Qy 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337  
Db 1176 AGAGGAGCAGCCTATGAATCTTCAACATATTTGATAATTAAGCCAGTATTCAGAGCTTC 1235  
Qy 338 SerThrAlaGlyTyrLysProGlySerIleGluGlyThrValGluPheLysAsnValSer 357  
Db 1236 TCAAGAAATGGGTACAACACGAGCAATTAAGAGAAATTTGAAATTTCAAAATATTCAC 1295  
Qy 358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377  
Db 1296 TTCAGTTACCCATCTCGAAAAGCGTTTCAGATCTTGAAGGCGCTCAACCTGAAGGTGCG 1355  
Qy 378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397  
Db 1356 AGCGGACAGACAGTGGCGCTGTTGGCAACAGTGGCTGTGGGAAACACCACTGTCCAG 1415  
Qy 398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417  
Db 1416 CTGCTGAGAGGCTCTACAGCCCAAGAGGCGTGTGAGTATCGACGAGGACATC 1475  
Qy 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437  
Db 1476 AGGACCATCAATGTGAGGTATCTCGGGAAATCATTTGGGTGCTGTGAGTACAGAACCTGTG 1535  
Qy 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457  
Db 1536 TTGTTTCCCACTATCGCTGAAACATTCGCTATGCGCGAGAAATGTCATCTATGAGC 1595  
Qy 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477  
Db 1596 GAGATCGAGAAAGCTGTCAAGAAAGCCAAATGCCCTATGACTTCATGAAACTGCCCAT 1655  
Qy 478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497  
Db 1656 AAATTTGCACATCTTGTGTGAGAGAGGACACAGCTGAGTGAGGACAGAAACACAGA 1715  
Qy 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517  
Db 1716 ATCGCCATTTGCTGCGCTGCTCGGCAACCCCAAGATCTCTTTGTTGGATGAGCGACA 1775  
Qy 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLys 537  
Db 1776 TCAGCCTTGGACAGAAAGTGAAGCGGTGTTGAGGTCTGAGGTCTGAGTAAAGGTAGAGAA 1835  
Qy 538 GlyArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557  
Db 1836 GCGCGGACTACCATTTGATAGACCCACCGCTTGTCTACAGTTCCAAATGCTGACATCATA 1895  
Qy 558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577  
Db 1896 GCTGGGTTTGATGTGTGTCTGAGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAG 1955

Qy 578 ArgGlyLeuTyrTyrSerLeuValMetSerGln----- 598  
Db 1956 AAGGGCATTTACTTCAAACTTGTTCATGACACAGCAGGAAATGAAATTTGAATTAGGA 2015  
Qy 589 ---AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 607  
Db 2016 AATGAAGTTGGTGAAGTAAATGAAATTTGATTAATTTAGACATGTCTTCAAAAGATTCA 2075  
Qy 608 ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625  
Db 2076 GCATCCAGTCTATCAGAGGAGATCACTCGCAAGATATCCGTGGACCAATGACCAA 2135  
Qy 626 LysAlaGluGluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeu 644  
Db 2136 GACAGGAAGCTTAGTACCAAGAGCGCTTGGATGAAGATGTACTCCAATTTCTTTGG 2195  
Qy 645 LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer 664  
Db 2196 CGTATCTAAAGTTGAATTCATCTGAATGGCCATATTTTGTGGTGTATTTCTGTGCC 2255  
Qy 665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684  
Db 2256 ATAGTAATGGAGCCTTGCACACGACATCTCAATAATATTTCTAAGGTGTAGGGGT 2315  
Qy 685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703  
Db 2316 TTCACAAGAAATACTGATGATGAACCAACGACATGATAGCAACTTGTTCCTTATTG 2375  
Qy 704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723  
Db 2376 TTTCTGATCCTGGGTCATTTCTTTTATTACATTTTCTTCTCAGGGCTTCACATTTGCC 2435  
Qy 724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743  
Db 2436 AAAGCTGGAGAGATCTCACCAAGCGACTCCGATACATGTTTTCAAATCCATGCTGAGA 2495  
Qy 744 GlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763  
Db 2496 CAGGATGTGAGCTGGTGTGATTAACCCCTAAAAACACCACTGGACATTTGACCACGCTC 2555  
Qy 764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783  
Db 2556 GCCAAGCATGCTGTCTCAAGTTAAAGGGGCTACAGGAGCCAGACTTGTCTCATCCAG 2615  
Qy 784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThr 803  
Db 2616 AATATGCAAAATCTTGGACAGGAATCATCATATCCCTATCTATGCTGGCAGTTGACA 2675  
Qy 804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823  
Db 2676 CTTCTACTCTTAGCAATTTGCCCATCATTTGCAATAGCAGAGTGGTTGAGATGAAATG 2735  
Qy 824 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 843  
Db 2736 TTGCTGTGGAACAGCACTTAAAGATAAGAGAGGCTAGAGGTTCTGGGAAGATGCTACT 2795  
Qy 844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863  
Db 2796 GAAGCAATAGAGAACTTCGCACTGTCTCTTTGACTCGGAGCAGAAAGTTTGAANAAT 2855  
Qy 864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883  
Db 2856 ATGTATGCCAGAGCTTGCAGATACCATCAGAAATGCTCTGGAAGAGACACAGCTTTT 2915  
Qy 884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903  
Db 2916 GGGATCACATTCCTTCCACCGAGGCATGATGATTAATTTTTCGTATGCTGCTTGTTCGG 2975  
Qy 904 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923  
Db 2976 TTTGGTCTTACTTGGTGGCAGCGGAATTAATGACATTTGAAATGTTCTTAGTATTTC 3035

Qy 924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943  
Db 3036 TCAGCTATTGCTTGTGGTCATGGCAGTGTGGCAGGTTCATTGCTCGTCACT 3095  
Qy 944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuLysLysProAsnIle 963  
Db 3096 GCCAAGCCAAAGTGTGGCATCGCATATCATCATCATGTAAGAAAGTCCCTCCATT 3155  
Qy 964 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983  
Db 3156 GACAGCTACAGCAGCCGAGCCCTGAAGCCTATACATTGGAAGGAATGTGAATTAAT 3215  
Qy 984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003  
Db 3216 GAAGTCGTGTCAACTACCCACCGCCCTGCATCCAGTGTTCAGGGGCTGAACCTG 3275  
Qy 1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023  
Db 3276 GAGTGAAGAAGGCCAGACGCTGGCCCTGGTGGGAGCAGTGGCTGCGGGAAGACACA 3335  
Qy 1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043  
Db 3336 GTGTCCTCAGCTGCTCAGCGGTTCTACGACCCCATCGCCGGAACAGTGTTCATGATGC 3395  
Qy 1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063  
Db 3396 AAAGAAGTAACCAACTAAATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGCCCAA 3455  
Qy 1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083  
Db 3456 GAGCCCATCTGTTGACTGCAGCATCGCTGGAACATCGCTTATGGAGACAAACGCCGG 3515  
Qy 1084 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaIleHisSerPheIle 1103  
Db 3516 GTGCTGTCAGCATGAGATCGAGAGCGGCGGCGAAGGACCAATCCACCATTCATT 3575  
Qy 1104 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123  
Db 3576 GAGTCACTGCCTGACAAATACACACACAGTAGGAGACAAAGGAGTCACTGCTCGGCT 3635  
Qy 1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 1143  
Db 3636 GGGCAGAAACACGCATCGCATAGCAGCGCCCTCGTCAGACAGCCTCACATTTTGCTC 3695  
Qy 1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163  
Db 3696 CTGGATGAAGCAACATCAGCCCTGGATACAGAGATGAAGGTTGTCCAGGAGCTCTG 3755  
Qy 1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183  
Db 3756 GACAAAGCCAGAGAGCGCGCACCTGATTTGATCGCTCACCGCTGTCCACCATCCAG 3815  
Qy 1184 AsnAlaAspIleIleValValLeuHisAsnGlyValIleLysGluGlnGlyThrHisGln 1203  
Db 3816 AACCGAGACTGATCGTGGTATTCAGATGGCAAGTGGCAAGGATGAGCAGATGCCACCCAG 3875  
Qy 1204 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
Db 3876 CAGCTGCTGCACAGAAAGCATATTTCTCCATGCTAGTGTGAGGCT 3926  
RESULT 8  
LOCUS AX105082 4279 bp DNA linear PAT 30-APR-2001  
DEFINITION Sequence 26 from Patent WO0123540.  
ACCESSION AX105082  
VERSION AX105082.1 GI:13921232  
KEYWORDS  
SOURCE dog.  
ORGANISM  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE 1 (bases 1 to 4279)  
AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and

Patten,C.J.  
P-glycoproteins and uses thereof  
Patent: WO 0123540-A 26 05-APR-2001;  
GENTEST CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1..4279  
/organism="Canis familiaris"  
/db\_xref="taxon:9615"  
17..3862  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC37723.1"  
/db\_xref="GI:13921233"  
/translation="MDPEGKGSABKNFWMGKSKKKEKKKPTVSTFAMFRYSN  
WLDRLYMLVGTMAAI IHGAALPLMLVFGNMTDSFANAGI SRNKTFPVI INESITNT  
QHFNHLEEMTTVAYYSGVAGVLAAYIVQSFWCLAGRAQLIKRQFQKFMRLRQ  
EIGFDVHDVGLNTRLTDDYSKNEGIDKIGMFFQSIATFTFGFI VGFTRGWLTL  
VILAI SPVLGSAI WAKILSSFTDKELLAYAKAGAAVEEVLAI RTVIAEGGQKEL  
EYNNKLEAKGIGIKKAITANISIGAAFLLIYASIALAFWGTSLVSSEYTGQVL  
TVFVSLIGAFSIGQASPISEAFANARGAAEYI FKIDNKPSIDSYSKSGHPDNIG  
NLEFNHVSYPFSREVKILKGLNFKVSGQTVALVNGSGCKSTTVQLMQRLYDPTD  
GMVICGQDIRTNVRLREITGVVQSPVL PATTI AENIRYGRNVTMDIEKAVKE  
ANAYDFIMKLPNKEDTLVGERGAQLSGGQKORIAIARALVRNPKILLDEATSAUDE  
SEAVQVALDKARKGRTTIVIAHRLSTVRNADVTAGPDGVLVEKGNHDELMKEGII  
FKLVMTQTRNEIELENAITGESKESDALESPDGSGLIKRRSTRSIHAPQGDQR  
KLGTEDLNENVPVFWRLIKLNSTEWPFYFVGIFCAIINGGLOPAPSIIPSRIGI  
PTRDEDPETKQNSMFSVLVLGIISIFITFLQGTFGKAGEILTKRLRYMFRSM  
LRQVSWDPDKNTTALTTLANDAAOVKAIGSLAVITQNIANLGTIIISIIYR  
WOLTLALLAIPII AIAAGVEMKLSGOALKDKKELEGAKIATEIENFTVVSILTR  
EQKEFYMAQSLQVPRYNSLRKAHIFGVSPSIQNMFFSYAGCFRGAYLVANFPMN  
FQDVLVFSALVFGMAVQVSSPADYAKAVSAHVIMII EKSPLDISYSYPHGLKP  
NTLEGNVTNFWNFYTRPDI PVLQSLVLEKKGQTLALVSGSCGKSTVVQLLERF  
YDLAQSVLIDGKEIKHLNWLRAHLGIVSQEPLFDPCSAENIAYGDNRSRVSHEE  
IVQAAKANIHHFTETLPEKYNTRVGDGTOLSGGQKORIAIARALVRPHILLDEA  
TSALDTESEKVVQBALDKAREGRTICIVIAHRLSTIQNALDLVVFQNGKVEHGTQQOL  
LAQKGFMSVMVQGAQR"  
BASE COUNT 1296 a 833 c 1009 g 1141 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,8e-227 Length: 4279  
Score: 3533.50 Matches: 683  
Percent Similarity: 73.13% Conservative: 237  
Best Local Similarity: 54.29% Mismatches: 269  
Query Match: 57.37% Indels: 69  
DB: 6 Gaps: 9  
US-09-873-409-5 (1-1222) x AX105082 (1-4279)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 167 ATGTTGGTGGGACAAATGGCTGCCATCATCCATGGAGCTGCACCTCCCTCTCATGATGCTG 226  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 227 GTTTTGGAAACATGACAGATAGCTTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 286  
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 287 CCAGTTATAATTAATGAAGAGTATTAGCAACAATACACAATTTTCATCAACCATCTGGAG 346  
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
Db 347 GAGGAATGACCACGATGCTTATTATACAGTGGGATCGGTGCTGCTGGTGGTGGCT 406  
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
Db 407 GCTTACATCCAGGTTTCACTTCTGGTCCCTGGCAGCAGGAGACAGATACTCAAAATTAGA 466  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99  
Db 467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATGGCTGGTTTACGCTGATGACGCT 526

QY 100 GlycylLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
DB 527 GGGAGCTTAACACCGGCTCACAGACGATGTCTCCAAAATCAATGAAGAAATGGCGAC 586  
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
DB 587 AAAATTGGAAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646  
QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
DB 647 ACAGTGGTTGGAGACTAACCTTGTGATTTTGGCCATCAGCCCTGTCTTGACATTTCA 706  
QY 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSer 178  
DB 707 GCCGCCATCTGGCAAGATACTATCTTCAATTAAGAACTCTTGGCCTATGCA 766  
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
DB 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTTGCCCTTTGGA 826  
QY 199 AlaGlnGlyLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218  
DB 827 GGCAAAAGAAAGAACTTGAA----- 847  
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeu 238  
DB 848 -----AGGTACAACAATAATTTA 865  
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
DB 866 GAAGAAGCTAAAGGAATTTGGGATAAAGAAAGCTATCAAGCCCAACATTTCTATTGTGGCC 925  
QY 259 ValTyrPhePheMetAsnGlyTrpTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278  
DB 926 GCTTCTTATGATCATGATCATATGCTCTGGCTTCTGGTATGGGACTCTTGGTC 985  
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
DB 986 CTCTCCAGTGAA-----TATACTATTGGACAAGTACTCACTGCTCTCTTTCTGTATTA 1039  
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
DB 1040 ATTTGGGGCTTTTATGATTTGGACAGGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099  
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
DB 1100 GGAGCCTCTATGAAATCTTCAAGATTAATGACATTAACCAAGCATTGACGCTATTCG 1159  
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358  
DB 1160 AAGAGTGGACATAAACCCAGATAATATTAAAGGAAATTTGGAAATTCAAAATGTTCACTTC 1219  
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
DB 1220 AGTTACCTCTTCGAAAAGAAGTTTAAGATCTTAAAGGCTCTCAACCTGGAAGGTTACAGAT 1279  
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
DB 1280 GGGCAGACAGTGGCGCTGGTTGGAAACAGTGGCTGGGGAAGACACACCGTGCAGCTG 1339  
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
DB 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGCTGTATTGATGACAGGACATTTAGG 1399  
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
DB 1400 ACCATAAATGTAAGGCATCTCTCGGAAATTTACTGGTGTGGTGTGAGTCAGAGCCTGTG 1459  
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458  
DB 1460 TTTCACCAACAGATGATGAAAACATTTGCTATGGCCCGGAAATGTCACCATGATGAG 1519  
QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478

DB 1520 ATTGAGAAAGCTGTTAAGGAAGCCCAATGCTATGATTTTATCATGAACACTACCTAATAAA 1579  
QY 479 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
DB 1580 TTTGACACTCTGTTGGAGAGAGAGGGCCGACGTGAGTGGTGGACAGAAACAGAAATC 1639  
QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
DB 1640 GCCATTGCTCGGCCCTGTTTCGCAACCCCAAGATTTCTTCTGCTGGATGAGGCAACGTCA 1699  
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538  
DB 1700 GCTCTGCACACTGAAGAGTGAAGCAGTGGTTTCAGGTGGCCCTGATAGGCCAGAAAGGC 1759  
QY 539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558  
DB 1760 CGGACTACCATTTGTGATAGTCTATGCTTGTACAGTTTCGTAATCCGATGTCATGCT 1819  
QY 559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578  
DB 1820 GGTTTTGATGATGGAGTCAATTGTGGAGAAAGAAATCATGATGAACCTCATGAAGAGAAG 1879  
QY 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598  
DB 1880 GGCATTACTTCAAACTTGTCAATGTCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936  
QY 599 SerMetTyrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
DB 1937 ATGCCACTGTGTAATCCAAAAGTGAAGTGATGCTTGAAATG---TCTCCAAAGAT 1993  
QY 619 IleLysSerAspPheIle----- 624  
DB 1994 TCAGGCTCCAGTTTATAAAGAAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053  
QY 625 ---AspLysAlaGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643  
DB 2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTCAATGAGAATGTACTCCAGTTTCTTTC 2113  
QY 644 LeuLysIleLeuLysAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663  
DB 2114 TGGAGATTTCTGAAGCTGAATCACTGAATGGCTTATTTTGTGGTGGTATATTTGT 2173  
QY 664 SerValLeuAsnGlyThrValHisProValPheSerIlePheAlaLysIleIleThr 683  
DB 2174 GCTATTATAACCGAGGCTGCACACCGACTTTTCAATAATATTTTCAAGGATTAAGGG 2233  
QY 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702  
DB 2234 ATCTTTACCCGAGATGAGGATCTCGAAACAAAACGACAGAAATAGTAACATGTTTCTGTA 2293  
QY 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722  
DB 2294 TTGTTTCTAGTCTCTGGAATATTTCTTTTATPACATTTTCTCCAGGCTTCACATTT 2353  
QY 723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742  
DB 2354 GGCAAGCTGGGAGATCTCACTAAGCGCTTCGATACATGCTGTTTTCAGATCCATGCTG 2413  
QY 743 TyrGluAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyLeuThrThrIle 762  
DB 2414 AGACAGGATGTACGTGTTGATGACCCCTAAACACCAACCATGAGCATTTGACAAACGAG 2473  
QY 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782  
DB 2474 CTTCGAATGATCGGCTCAAGTTAAAGGGCTATAGTTCAGGCTTGTCTGTCATACC 2533  
QY 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802  
DB 2534 CAGAAATAGCAAACTCTGGACAGGCAATTTATATCTTATCTTATCTATGTTGGCAATTA 2593  
QY 803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822







Db 250 ATGGTGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSer-----GlyCys 33  
Db 310 GTGTTTGGAGACATGACGGATACCTTTGCAAAATGCGAGAAATTTAGGAGATTTAGGAGCT 369  
Qy 34 LeuValGlnThrAsnThrTySer----- 41  
Db 370 CTGTTGTTTAAACACACTAATAGCAGTAATATCACTGATACAGTGCCTCGTCATGAATCTG 429  
Qy 42 -----PhePheArgLeuThrLeuTyTyValGlyIleGlyValAlaLeuLeu 58  
Db 430 GAGGAAGATATGACCATGCTTACAGTATGCTTATTAATGCTGGAATTTGGTGGGCTGGTT 489  
Qy 59 PheGlyTyIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrTyLeuArgIle 78  
Db 490 GCTGCTTACATTCAGGTTTCATTTTGGTGGCTGGACGCTGGAAGACAAATACACAAAT 549  
Qy 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98  
Db 550 AGAAAAACAGTTTTTTCATGCTAATCGACAGGATAGCTGGTTTGAATGTCACGAT 609  
Qy 99 IleGlyGluLeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGly 117  
Db 610 GTTGGGAGCTTAAACCCCGCTTACAGATGATGCTCTCAAGATTAATGAAGAAATGGT 669  
Qy 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137  
Db 670 GACAAAATTTGAATGTTCTTTCAGTCAATGGCAACATTTTCACTGGGTTTATAGTAGGA 729  
Qy 138 LeuValLysGlyTyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157  
Db 730 TTTACAGCTGGTGGAGACTAACCTTGTGATTTGGCCATCAGTCTCTGTTCTGGACTG 789  
Qy 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyI 177  
Db 790 TCAGCTGCAGTCTGGCAAGATACTGCTTCATTACTGATAAAGAACTCTTAGCTTAT 849  
Qy 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197  
Db 850 GCAAAAGCTGAGCAGTAGCTGGAAGAGTCTTGGCAGCAATTAGAATGATGATGCAAT 909  
Qy 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrp 217  
Db 910 GGAGGACAAAAGAAAGAACTCGAA----- 933  
Qy 218 PheTyIleProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyThrGlnAsn 237  
Db 934 -----AGGTACAAACAAAAT 948  
Qy 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257  
Db 949 TTAGAAGAAGCTTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAAATTTCTATAGGT 1008  
Qy 258 AlaValTyPhePheMetAsnGlnTyTyTyGlyLeuAlaPheTrpTyGlyThrSerLeu 277  
Db 1009 GCTGCTTCCTGCTTACTATGCTATGCTATGCTCTGGCCCTCTGGTATGGACACCTG 1068  
Qy 278 IleLeuAsnGlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerVal 297  
Db 1069 GTCCTCTCAAAGGAA-----TATTCTATTGGACAAGTACTCACTGATTTCTTTCTGTA 1122  
Qy 298 IleHisSerSerTyIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317  
Db 1123 TTAATTTGGGCTTTTAGTGTGGACAGGCACTCTCAAGCATTTGAAGCATTTGCCAAATGCA 1182  
Qy 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337  
Db 1183 AGAGGAGCAGCTTTTGAATCTTCAAGATATTTGATAATAGCCCAAGTATTGACAGCTAT 1242  
Qy 338 SerThrAlaGlyTyIleProGluSerIleGlyThrValGluPheLysAsnValSer 357  
Db 1243 TCGAAGAGTGGGCACAAACCAAGATAATTAAGGGAAATTTGGAATTCAGAAATGTTTCAC 1302

Qy 358 PheAsnTyIleProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377  
Db 1303 TTCAGTTACCATCTCGAAAAGAGTTAAGATCTTGAAGGGCTGAACCTGTAAGGTGCAG 1362  
Qy 378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397  
Db 1363 AGTGGGCAGACGTCCTGGTGTGAAACAGCGGCTGTGGGAAGAGACAAACGTCCTCAG 1422  
Qy 398 LeuLeuGlnArgLeuTyIleAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417  
Db 1423 CTGATCGACAGGCTTTATGACCCACAGAGGGCATGTCAGTGTGTTGATGACAGGATATT 1482  
Qy 418 ArgAlaLeuAsnValArgHisTyIleArgAspHisIleGlyValValSerGlnGluProVal 437  
Db 1483 AGGACCATTAACGTAAGTTTCTACGGGAAATCATCGGTGTGTTGAGTCAGGAACCTGTA 1542  
Qy 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyIleArgAspAspValThrAspGlu 457  
Db 1543 TTTGTTGGCCACCACTAGCTGAAACATTCGCTATGCTGCTGTAAGATGTCACCATGGAT 1602  
Qy 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyIleAspPheIleMetGluPheProAsn 477  
Db 1603 GAGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGACTTTATCATGAAATGCTCCTCAG 1662  
Qy 478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497  
Db 1663 AATTTTGACACCTGTTGGAGAGAGGGGCCACGTCAGTGTGGGCAAGAGCAGAGG 1722  
Qy 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517  
Db 1723 ATCGCATTCGACGTCCTGTTGCGCAACCCCAAGATCTCTCTGTCGACGAGGCCACG 1782  
Qy 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLys 537  
Db 1783 TCAGCTTGGACACAGAAAGTGAAGCAGTGTTTCAGGTGCTCTGGATAGGCCAGAAAA 1842  
Qy 538 GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557  
Db 1843 GGTTCGGACCACTTGTGATAGCTCATCTGTTCTACGGTTCGTAATCCGACGTCATC 1902  
Qy 558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577  
Db 1903 GCTGGTTTCGATGATGAGTCAATTTGGAGAAAGAAATCATGATGAGCTCATGAAGAG 1962  
Qy 578 ArgGlyLeuTyIleSerLeuValMetSerGlnAspIle----- 590  
Db 1963 AAGGCAATTTACTTTCAAACTTGTCAACATGCACAGCAGCAGGAAATGAAATGATAGAA 2022  
Qy 591 LysLysAlaAspGlu-----GlnMetGluSerMetThrTyIleSerThrGluArgLys 607  
Db 2023 AATGCAGCTGATCAATCCAAAGATGAAATGTATACCTTGGAAATGTCTTCACATGATTCA 2082  
Qy 608 ThrAsnSerLeu-----ProLeuHisSerValLysSerIleLysSerAsp 622  
Db 2083 GGATCCAGCTTAATAAGAAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA 2142  
Qy 623 PheIleAspLysAlaGluGluSerThrGlnSerLysGluLysSerLeuProGluValSer 642  
Db 2143 -----GACAGAAGCTTAGTACCAAGAGGCTCTGGATGAAAGATATACCTCCAGTTTC 2196  
Qy 643 LeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeu 662  
Db 2197 TTTTGGAGGATTATGAAGCTAATTTAACTGAGTGGCCTTATTTGTTGTTGTTGTTATTT 2256  
Qy 663 AlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIle 682  
Db 2257 TGTGCCATTAATAATGGAGTCTGCAACCCAGCATTTGTCAGTAATATTTTCAAGATTATA 2316  
Qy 683 ThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTySer 701  
Db 2317 GGGATTTTCAAGAAATGATGATGCGCAACCAACCAAGACAGATAGTAGTAATGTTTCA 2376

QY 702 MetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPhe 721  
Db 2377 CTATTTCTTCTAGCTCTCGAATTTGTTCTTTATTACATTTTCTCAGGGCTTCACA 2436  
QY 722 TyrGlyArgAlaGlyLeuIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMet 741  
Db 2437 TTTGGCAAGCTGGAGAGATCTCACAGCGGCTCGATACATGGTTTTCATCCATG 2496  
QY 742 LeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThr 761  
Db 2497 CTCAGACAGATGTGAGCTGTTTGTATGACCTCAAAACACCACTGGAGCATTGACTACC 2556  
QY 762 IleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeu 781  
Db 2557 AGGCTGCCAATATGCTGCTCAAGTTAAAGGGCTATAGGTTCCAGGCTGTCTAAT 2616  
QY 782 ThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGlu 801  
Db 2617 ACCAGAATATACCAATCTTGGGACAGGATATATATCTTAATCTATGTTGGCAA 2676  
QY 802 MetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThr 821  
Db 2677 CTGACACTGTACTCTTAGCAATTTGTACCATCATTCGATAGCAGGAGTTGTGAATG 2736  
QY 822 AlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIle 841  
Db 2737 AAAATGTTGCTGGCAAGCACTGAAAGATAAGAAAGAACTAGAAAGTGTGGGAAGATC 2796  
QY 842 AlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPhe 861  
Db 2797 GCTACTGAAACATAGAAACTTTCCGAACCTGTTGTTCTTTGACTCAGGAGCAAGATT 2856  
QY 862 GluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGln 881  
Db 2857 GAACATATGATGATCAGAGTTTGAGGTACCATACATACAGAAACTCTTTGAGGAAGCACAC 2916  
QY 882 IleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGly 901  
Db 2917 ATCTTTGGAATCAGTTTCTTCCAGCAGGCAATGATGATTTTCTATGCTGATGT 2976  
QY 902 PheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIle 921  
Db 2977 TTTCCGTTTGGAGCCCTACTTGGTGGCACAPAGTCTCATGAGCTTTCAGGATGTTCTGTTA 3036  
QY 922 ValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaPro 941  
Db 3037 GTATTTTACAGTGTGCTTGTGTGCTGCGCGTGGGCAAGTCAGTCAATGTCCTCT 3096  
QY 942 GluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysPro 961  
Db 3097 GACTATGCCAAAGCCAAAGTATCAGCAGGCCACATCATCATGATCATTTGAAAAACCCCT 3156  
QY 962 AsnIleAspSerArgSerGlnGluGlyLysProAspThrCysGluGlyAsnLeuGlu 981  
Db 3157 TTGATTTGACAGCTACAGCACAGAAGGCTTAAGCCGGAACATTTGGAAGAAATGTCA 3216  
QY 982 PheArgGlnValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeu 1001  
Db 3217 TTTAATGAAGTTGTTAATCACTATCCACCCGATGTCATCCAGTCTTCAGGGGCTG 3276  
QY 1002 SerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyLysLys 1021  
Db 3277 AGCTGGAAGTGAAGAGGCCAGACGCTGCTGCTGGGCGAGCAGTGCCTGTGGGAAG 3336  
QY 1022 SerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPhe 1041  
Db 3337 AGCAGCGTGTCTCAGCTCTCTGGAGCGGTTCTATGACCCCTTGGCGGGAAGTGTCTT 3396  
QY 1042 AspGlyValAspAlaLysGluLeuAsnValClnTrpLeuArgSerGlnIleAlaIleVal 1061  
Db 3397 GACGGCAAGAAATAAAGCAACTGAATGTTTCAGTGGCTCCGAGCACACCTGGGCGATG 3456  
QY 1062 ProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsn 1081

Db 3457 TCCAGGAGCCCATCTGTTTGACTGTCAGCATTAGTGAAACATTGCCCTATGAGACAAC 3516  
QY 1082 SerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 1101  
Db 3517 AGCCGGTGTCTCAGAGGAAGATCGTGAGGCGAGCCCAAGGAGCCATATACAGCC 3576  
QY 1102 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 1121  
Db 3577 TTTATCGAGTCACTGCTTAATAATATAGCACAGATAGGAGCAAAAGAACTCAGCTC 3636  
QY 1122 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLysProLysIle 1141  
Db 3637 TCTGGTGGCCAGAAACACGCAATGTCATAGCTCGGCCCTTGTAGACAGCCTCATATT 3696  
QY 1142 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 1161  
Db 3697 TTGCTTTGGATGAGCCACATCAGCTCTGGATACAGAAAGTGAAAGGTGTTCAGAA 3756  
QY 1162 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 1181  
Db 3757 GGCCTGGACAAAGCCAGAGAGGCGGTACCTGCTGATTGCTCACCCTGTCCACC 3816  
QY 1182 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 1201  
Db 3817 ATCCGAATGCACTTAATAGTGTGTTTCAGATGGCAGAGTCAAGGAGCACGSCACA 3876  
QY 1202 HisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValValAlaGlnSer 1220  
Db 3877 CATCAGCAGCTGCTGGCACAGAAAGCATCTATTTTCATGGTCAAGTGTCCAGGCT 3933  
RESULT 10  
AF257746 4927 bp mRNA linear ROD 09-MAY-2000  
LOCUS Rattus norvegicus multidrug resistance protein 1a (Pgyl) mRNA,  
DEFINITION complete cds.  
ACCESSION AF257746  
VERSION AF257746.1 GI:7739772  
KEYWORDS  
SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 4927)  
AUTHORS Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,  
Meijer,D.K.F. and Muller,M.  
TITLE Cloning and functional characterization of the rat multidrug  
resistance protein Mr1a  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 4927)  
AUTHORS Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,  
Meijer,D.K.F. and Muller,M.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2000) Div. Gastroenterology & Hepatology,  
University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,  
The Netherlands  
FEATURES  
source Location/Qualifiers  
1. .4927  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/sex="male"  
/cell\_type="hepatocyte"  
/tissue\_type="liver"  
/dev\_stage="adult"  
1. .4927  
/gene="Pgyl"  
85.3904  
/gene="Pgyl"  
/note="ATP-binding cassette transport protein; Abcb1a"  
/codon\_start=1  
/product="multidrug resistance protein 1a"



```
QY 526 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
DB 1760 GCGTGGTTTCAGGCGCTCTGGATAGGCTAGAGAGCGCGACCATTTGTAGTCT 1819
QY 546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
DB 1820 CACGCTTGTTACAGTTCCGAATCTGACGTCAATGCTGGTTTGTATGGTGTGCAAT 1879
QY 566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
DB 1880 GTGAGCAAGAAATCATGATGAGTCATGAGAGAGAAAGAAATTTACTTCAAATGTGC 1939
QY 586 MetSerGln-----AspIleLysAlaAspGluGlnMetGluSer----- 599
DB 1940 ATCATCAGACAGCAGGAAATGAATTTGAATTTAGAAATGAAGCTTGTGAATCTAAGAT 1999
QY 600 -----MetThrTyrSerThrGluArgLysThrAsnSerLeu-----ProLeu 613
DB 2000 GGAATTGATATGTGGACATGCTCTCAAAAGATTTCAGGATCCAGTCTAATAAGAGAGA 2059
QY 614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632
DB 2060 TCAACTCGCAAAAGCATCCGTGGCGGCACATGATCAAGACCGGGAACTTAGCACCAAGAG 2119
QY 633 SerLysGluLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 652
DB 2120 GCTCTGATGACGACGATACCTCCAGCTTCCTTTTGGCGGATCTCTGAAGTTGAATCAA 2179
QY 653 GluTrpProPheValLeuGluGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
DB 2180 GAATGGCTTATTTTGGTGTGGTATTTTGGCCATAATAATGAGGCTTGCACCA 2239
QY 673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 692
DB 2240 GCATCTCCATAATATTTCAAGAGTTGTAGGGGTTTTACAAAAATGACACCCCTGAA 2299
QY 693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
DB 2300 ATCCAGCGGAGAACACCAACTCTTCAATCCATGCTGACAGACAGACATAGCTGGTTGATG 2359
QY 712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731
DB 2360 TTCAATACGTTTTCCTTCAGGGCTTCACATTTGGCAAGCTGGAGAGATCCCTCACCAAG 2419
QY 732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
DB 2420 CGACTCGATACATGCTCTCAATCCATGCTGACAGACAGACATAGCTGGTTGATGAC 2479
QY 752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
DB 2480 CCTAAAACACCAACAGGAGCGCTGACCACAGGCTTGCCCAATGACGCTGCTCAAGTGAAA 2539
QY 772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
DB 2540 GGGGCTACAGGGCTAGGCTTGCTGTTATTATCCAGAACATACAAATCTTGGGACAGGC 2599
QY 792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
DB 2600 ATCATCATATCCCTGATCTACGCTGCAATGACATTTTACTCTTAGCAATTTGTCCC 2659
QY 812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831
DB 2660 ATCATTTGCTATAGCAGGAGTGGTGAATTTGAATTTGTCTGGACAAAGCTGAAAGAT 2719
QY 832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
DB 2720 AAGAAGGAATAGAGGTTCTGGGAAGATCGCTACAGAAAGCAATTTGAATTTTCCGCACT 2779
QY 852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
DB 2780 GTGCTCTCTTGTAGCTCGGAGCAGAGATTTGAATACTATGATGCCACAGAGCTTGCAGATA 2839
QY 872 GlnHisArgAsnThrSerLysAlaGlnIleIleGlySerCystyrAlaPheSerHis 891
```

```
DB 2840 CCATACAGAAATGCTTTGAAGAAAGCGCGCTTTTGGGATCACATTTCCTTCACCCAG 2899
QY 892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
DB 2900 GCCATGATGATTTCTCCTATGCTGCTGTTTCGGGTTTGGTGCCTACTTGGTGGCAGCA 2959
QY 912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
DB 2960 GAACCTCATGACATTTGAAAATGTTCTGTAGTATTCTCAGCTATTGCTTGTGGTGCATG 3019
QY 932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
DB 3020 GCAGTGGGCGAGCTCAGTTCAATTCGCTCTGACTACGCGAAAAGCCAAAGTCTCGGCATCC 3079
QY 952 HisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
DB 3080 CACATCATCAGGATCATTTGAGAAATCCCTGAGATTGACAGCTACAGCAGGAGGCTTG 3139
QY 972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
DB 3140 AGCCCTAATATGTTGAAGAAATGTGAATTTAATGGAGTCATGTTCAACTATCCACC 3199
QY 992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
DB 3200 CGACCCCAACATCCAGTGTCTCAGGGGCTGAGCCTAGAGGTGAAGAAAGCGCAGCGCTG 3259
QY 1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeu 1031
DB 3260 GCCTCGTGGGAGCAGTGGCTGCGGAGAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTC 3319
QY 1032 TyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
DB 3320 TATGACCCCATGCCGGAACAGTGTCTTAGATGGCAAGAAATAAAGCAACTCAATGTC 3379
QY 1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
DB 3380 CAGTGGCTCGCGCCACCTCGGCATTTGTCTCCAGAGGCCATCTCTGTTGACTGCAGC 3439
QY 1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
DB 3440 ATCCGCGAGAACATTCCTACGGAGACACAGCCGTGCTGTCTCATGAGGAGATCGTG 3499
QY 1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
DB 3500 AAGGCAGCCAAAGGAGCCCAACATCCACAGTTCATCGACTCACTGCTGAGAAATACAAC 3559
QY 1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
DB 3560 ACCAGAGTGGGAGACAAAGGAGCTCAGCTGTCTGGGCGGCGAGAGCGCATCGGCATC 3619
QY 1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeu 1151
DB 3620 GCGCGGCTCTGTACAGACGCTCACTCTTACTTCTGTGATGAAGCAGCATCAGCTCTG 3679
QY 1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171
DB 3680 GATACGGAGAGTGAAAGGTGCTCCAGGAAGCGCTGCAAAAGCCAGGAGGCCGCCACC 3739
QY 1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeu 1191
DB 3740 TGCATTTGATCGCGCAGCGCTGTCCACCATCCAGAACGACACTTGATCGTGGTGATT 3799
QY 1192 HisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211
DB 3800 CAGACGGCAGCTCAAGGAGCAGCGCACCCACAGCAGCTGCTGCCCAAGAAAGGCATC 3859
QY 1212 TyrPheLysLeuValAsnAlaGlnSer 1220
DB 3860 TATTTCGATGCTGTCAGTGTGCAGGCT 3886
RESULT 11
AX322787
```

LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002  
DEFINITION Sequence 1 from Patent WO0192877.  
ACCESSION AX322787  
VERSION AX322787.1 GI:18093766  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Sorrentino, B. and Schuetz, J.  
TITLE Method of identifying and/or isolating stem cells  
JOURNAL Patent: WO 0192877-A 1 06-DEC-2001;  
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)  
FEATURES  
source  
1..3860  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1135 a 746 c 957 g 1022 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.9e-227 Length: 3860  
Score: 3530.50 Matches: 682  
Percent Similarity: 74.00% Conservative: 246  
Best Local Similarity: 54.39% Mismatches: 263  
Query Match: 57.32% Indels: 63  
DB: 6 Gaps: 10  
US-09-873-409-5 (1-1222) x AX322787 (1-3860)  
Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 151 ATGGTGGGAACCTTTGGCTGCCATCATCCATGGGCTGGACCTCTCTCATGATCGT 210  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAGATCTGATGCA 270  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 271 AACATCACTAATAGAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330  
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaLeuIlePheGlyTyr 61  
Db 331 ATGACCATATGCTTATTATACAGTGGAAATGGTCTGGGGTGGTGGTCTGCTTAC 390  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 391 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAAGACAAATACACAAAATTTAGAAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGTATATATGCGACAGGATAGCTGGTTGATGTCACGATGTTGGGGAG 510  
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 511 CTTAAACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTATTTGGTGAATAATT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GGTGGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCTGTTCTGGACATGTCAGCTGCT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 691 GTCTGGCAGAGATACACTCTTCATTACTGATAAAGACTCTTAGGATGATGCAAAAGCT 750  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCTGTAATGCTGACGTCATCGTGGTTTC 1803

Db 751 GGAGCAGTAGCTAGCGAAGAGGTCTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 810  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 811 AAGAAAGAACTTGAA-----AGGTACAAACAAAATTTAGAAAGAA 825  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 826 -----AGGTACAAACAAAATTTAGAAAGAA 849  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 850 GCTAAAAGAAATTTGGGATAAGAAAGCTATTACAGCAATAATTTCTATAGGTGCTGCTTC 909  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 910 CTGCTGATCTATGCACTCTTATGCTCTGGCTTCTGTTATGGGACCACTTGGTCTCTCA 969  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 970 GGGGAA-----TATCTATTGGCAAGTACTCATCTGTATTCTTTCTGTATTAAATTGGG 1023  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1024 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGCAAAATGCAAGAGAGCA 1083  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1084 GCTTATGAATCTTCAAGATATTGATTAAGCCAGATATTGACAGCTATTGCAAGAGT 1143  
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1144 GGGCACAAACCATATAATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203  
Qy 361 ProSerArgProSerIleLysLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1204 CCATCTCGAAAAGAACTTAAGATCTTGAAGGGCTGAACCTGAAGGTGCAAGATGGGCA 1263  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1264 ACCTGGCCCTGTTGGAAACAGTGGCTGGGAAGAGCAACAGTCCAGCTGATGTCAG 1323  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspLysAsnAspIleArgAlaLeu 420  
Db 1324 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGATGTCAGCAGGATATTAGGACCAT 1383  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1384 AATGTAAGSTTTCTACGGGAAATCATTTGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGCC 1443  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
Db 1444 ACCCATAGCTGAAAACCATTCGCTATGGCCGTAAGATGTCACCATGATGATGATGAG 1503  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1504 AAAGCTGTCAAGGAAGCAATATGCTATGACTTTATCATGAAGACTGCTCATATAATTGAC 1563  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1564 ACCCTGGTGGAGAGAGGGGCCCATGTTAGTGGTGGGAGAGAGAGATCGCCATT 1623  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1624 GCACGTGCCCTGGTTGCAACCCCAAGATCTCTCTGCTGATGAGGCCAGCTCAGCCTTG 1683  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 540  
Db 1684 GACACAGAAAGCAAGCAGTGGTTCAGTGGCTCTGGATAAGGCCAGAAAAGGTCGAGCC 1743  
Qy 541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560  
Db 1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCTGTAATGCTGACGTCATCGTGGTTTC 1803

Qy	561	LyAspGlyMetLeuAlaGluIysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu	580
Db	1804	GATGATGGAGTCATTGTGGAGAAAGAAATCATGATGAATCATCAAGACAGAAAGGCAATT	1863
Qy	581	TyrTyrSerLeuValMetSerGln-----AspIleLysLysAlaAsp	594
Db	1864	TACTTTCAAACCTGTGCAAATGCACAGACAGAGAAATGAAGTTGAAATAGAAAATGCAGCT	1923
Qy	595	GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis	614
Db	1924	GATGAATCCAAAGTCAAATTGATGCTTTGGAAATGCTCTCAAATGATTCAAGATCCAGT	1983
Qy	615	SerVal-----LysSerIleLys----SerAspPheIleAspLysAla	627
Db	1984	CTAATAGAAAAAGATCAACTCGTAGAGCTGTCGTGGATCACAAGCCCAAGACAGAAAG	2043
Qy	628	GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu	647
Db	2044	CTTAGTACCAAGAGGCTCTGGATGAAAGTAGTACTCTCAGTTTCTTTTGGAGGATTATG	2103
Qy	648	LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn	667
Db	2104	AAGCTAAATTTAACTCAATGGCCCTATTATTGTGTGGTATTATTTGTGCCATTATAAAT	2163
Qy	668	GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn	687
Db	2164	GGAGGCGCTGCAACAGCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTCAAGA	2223
Qy	688	--AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle	706
Db	2224	ATTGATGATCCTGAAACAAACACAGAGAATAGTAACTGTGTTTTCACTATGTTTCTAGCC	2283
Qy	707	LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly	726
Db	2284	CTTGGAAATTTCTTTTATTACATTTTTCCTCAAGGTTTCCACATTTGGCCAAAGCTGGA	2343
Qy	727	GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle	746
Db	2344	GAGATCCTCCACCAAGGGCTCCGATACATGGTTTTCGATCCATGCTCAGACAGGATGTG	2403
Qy	747	AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp	766
Db	2404	AGTTGGTTTGATGACCCCTAAACACACCATCGAGCAATTGACTACCGGCTCGCCAAATGAT	2463
Qy	767	IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr	786
Db	2464	GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTGTGCTGTAATACCAGAAATATAGCA	2523
Qy	787	AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle	806
Db	2524	AAATCTGGGACAGAAATAATATCCTTCATCTATGGTTGGCAACTCAACACTGTACTC	2583
Qy	807	LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly	826
Db	2584	TTAGCAATTGTAACCATCATTCGATAGCAGAGATTGTGAAATCAAAATGTTGCTCGGA	2643
Qy	827	PheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu	846
Db	2644	CAAGCACTGAAAGATAAGAAAGAACTAGAAAGGTGCTGGAGAGTCGCTACTGAAGCAATA	2703
Qy	847	GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu	866
Db	2704	GAAAACTTCGAACCGTGTCTTTTGACTCAGGACGAGAAGTTTGAACATATGTATGCT	2763
Qy	867	GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys	886
Db	2764	CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAGACACATCTTTTGGATTACA	2823
Qy	887	TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla	906
Db	2824	TTTTTCCTTACCCAGGCAATGATGTAATTTTCTATGCTGGATGTTTCCGGTTTGGAGCC	2883

Qy	907	TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle	926
Db	2884	TACTTGGTGGCACAATAAATCATAGAGCTTTGAGGATGTTCTGTGTAGTATATTTTACGCTGTT	2943
Qy	927	AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla	946
Db	2944	GTCTTTGGTGCCATGGCCGTGGGCAAGTCAGTTCATTGTTGCTCTGACTATGCCAAGGCC	3003
Qy	947	LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg	966
Db	3004	AAATATATCAGCAGCCACATCATATGATCATTTGAAAAAACCCTTTGATTGACAGCTAC	3063
Qy	967	SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuLeuGluPheArgGluValSer	986
Db	3064	AGCAGGAAGGCCATAATGCCGAACACATTTGGAAGAAATGTTCATATTTGGTGAATGTGA	3123
Qy	987	PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu	1006
Db	3124	TTCAACTATCCACCCAGCGAGACATCCAGATGCTTCAGGACCTGAGCGCTGAGGTGAAG	3183
Qy	1007	ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln	1026
Db	3184	AAGGCCCAGACGCTGGCTCTGTGTGGCAGCAGTGGCTCTGGGAAGAGCACAGTGGTCCAG	3243
Qy	1027	LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla	1046
Db	3244	CTCTGTGAGCGGTCTTACGACCCTTTGGCAGGGAAGTGCCTTTGATGGCAAGAATA	3303
Qy	1047	LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal	1066
Db	3304	AAGGACTGAATGTTCAAGTGGCTCCGAGCACAACCTGGGCATCTGTCTCCAGAGCCCATC	3363
Qy	1067	LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro	1086
Db	3364	CTGTTTGACTGCAGCATTCGTGAGAATTCCTATGAGACACAACAGCCGGGTGGTGCA	3423
Qy	1087	LeuAspGluIleLysGluAlaAlaAlaAlaAlaAsnIleHisSerPheIleGluGlyLeu	1106
Db	3424	CAGGAAGAGATCGTGAAGGCGCAAGAGGAGGCCAACATACATGCTTCATCGACTCATG	3483
Qy	1107	ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys	1126
Db	3484	CCTTAATAATATAGCACTAAAGTAGAGACAAGGAACCTCAGCTCTCTGGTGGCCAGNAA	3543
Qy	1127	GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu	1146
Db	3544	CAAGCATTCGCCATAGCTCGTCCCTTGTTAGACAGCCTCATATTTCTTTGGATGAA	3603
Qy	1147	AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla	1166
Db	3604	GCCACGTCAGCTCTGGATATCAGAAAGTGAAGAGTGTGTCGAAGAGCCCTCGCAAGAGCC	3663
Qy	1167	ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp	1186
Db	3664	AGAGAAGCCGCACTTCATTTGATGTCTACCCGCTGTCCACCATCCAGAATGCGAGAC	3723
Qy	1187	LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeu	1206
Db	3724	TTAATAGTGGTGTTCAGAAATGGCAGAGTCAGAGGAGCATGGCAGCATCAGCAGCTGCTG	3783
Qy	1207	ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	1220
Db	3784	GCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT	3825

RESULT 12

AX105078	LOCUS	AX105078	4279 bp
DEFINITION			
Sequence 22 from Patent WO0123540.			

AX105078
AX105078.1
GI:13921228

## KEYWORDS

SOURCE ORGANISM dog. Canis familiaris

# ORGANISM CAUSING TAILROT



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Reif,T.C. and  
Patten,C.J.  
P-glycoproteins and uses thereof  
Patent: WO 0123540-A 22 05-APR-2001;  
GENTEST CORPORATION (US)  
FEATURES  
source  
1. .4279  
/organism="Canis familiaris"  
/db\_xref="taxon:9615"  
17. .3862  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC37721.1"  
/db\_xref="GI:13921429"  
/translation="MDPEGRGSGAEKFNWKMKGSKKNEKKKPTVSTFAMFRYSN  
WLDRLVMLVGTWAAIITHGAALPLMLVFNMTDSFANAGISRNKTPFPVILINESITNT  
QHFINLEEMTAYYYGIGAGVLVAAYIOVSFMCCLAAGRIQILKIRKQFHAIMRQ  
EIGFEDVDGELNTRLTDVSKINIGIDKIMRFQSTATFTFTFVIGFTGWLTL  
VILAISPVGLSNAIWAKILLSFTDKELLAYAKAGAVEVLAAIRTVIAFGQKEL  
ERYNKLSEAKGIGIKKAITANISIGAFLIYASTALAFWTGTSILVLSSEYSGQVL  
TVFFSLIGAFSGQSPSIEAPANARGAAYEIFKIIDNKPSIDSYSGSKHPDNKIG  
NLEFNKHFYSYRKEVKILKGLNKVQSGQTVLVNCGSKSTTVQLMORLYDPTD  
GMVICDQDIRTINVRLREITGVSOEPLVATTIAENIRYGRNVNMTDEIKAVKE  
ANAYDPIMLKPNKFDLVGERGAOLSGOKORIAIARALVRNPKILLDEATSLDTE  
SEAVQVALDKARKRTIVIAHRLSTVRNADVIAGFDGVIIVEKGNHDELMEKGIY  
FKLVMTQRTNEIENATGESKSDALEMPSKGSLSIKRRSTRRSIHAPQODR  
KLGTEDLNENVPVPSFVFIKLNSTENFYFVVGIFCAIINGLOPAPFISPIRIGI  
FRDEDPYKQNSNMFVLVIGIISITFTFLOQTFKGAGELITKRLRYMVRSM  
LRDVSVDPPKNTGALTRLANDAAOVKGASRLAVITONIANLGTIISLIY  
WOLTLILAIVPIIATGVVEMKLSQALKKKELEGAGKATAEIENGRVTVSITR  
EQKFVYQAISLOVPRNSLRKAIHFQSFSTQAMMFYSIAGCFRFGAILVANBPMN  
FQDLVLSAIVFAGVAQVGSFADYAKAVSAAHVIMIEKSLSDISYSPHGLKP  
NTLEGNVTFENFVPTDPIVPLQGLSLEKVGQTALVSGSGSKSTVQVLLERF  
YDPLAGSVIDGKEIKHLNVQMLRAHLGIVSEKPIIDGCSIAENTAYGDSNRVSHEE  
IMQAAKEANHFIEITLPEKYNTRVGDKTQLSGGOKORIAIARALVROPHILLDEA  
TSALDPEYKVOEALDKAREGTCIVIAHRLSTIQNALDIVVFQNGKVEHGTHQOL  
LAQGGIYFNSVVOAGAKR"  
BASE COUNT 1295 a 833 c 1008 g 1143 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,45e-227 Length: 4279  
Score: 3530.50 Matches: 682  
Percent Similarity: 73.13% Conservative: 238  
Best Local Similarity: 54.21% Mismatches: 269  
Query Match: 57.32% Indels: 69  
DB: 6 Gaps: 9  
US-09-873-409-5 (1-1222) x AX105078 (1-4279)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 167 ATGTTGGTGGGACAAATGGCTGCCATCATCGAGCTGCACCTCCCTCTCATGATCGT 226  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGACAGAAATTTCAAGAAACAAACTTTT 286  
Qy 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44  
Db 287 CCAGTTATATTAATGAAGATTATGACGACATACACAACTTTTCATCACCATCTCGAG 346  
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePhe 59  
Db 347 GAGGAATGACCAGTATGCCTATTATTACAGTGGGATCGGTGCGTGGCTGGCT 406  
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
Db 407 GCTTACATCCAGGTTTCATTCTGGTGGCTGGCAGCAGGAAGACAGACTACTCAAAATTAGA 466

Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
Db 467 AAACAATTTTTTTCATGCTATCATGCAGCAGAGAGATTGGCTGGTTTTTCACGTGATCAGTT 526  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 527 GGGGAGCTTAAACACCGGCTCACAGCATGTCTCAAAATCAATGAAGAAATGGCGAC 586  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 587 AAAATTGGATGTTCTTTCAATCAATAGCAACATTTTACCAGTTTATAGTGGGGTTT 646  
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 647 ACACGTGGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTCA 706  
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 707 GCGGCCATCTGGGCAAGATATCTATCTTACTGTATAAAGAACTCTTGGGCTATGCA 766  
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 767 AAACGTGGAGCAGTAGCTGAGAGAGCTTTAGCAGCACTGTGATGCTGCTTGGGA 826  
Qy 199 AlaGlnGluLeuGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe 218  
Db 827 GGAACAAAAGAAAGAACTTGAA----- 847  
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeu 238  
Db 848 -----AGGTACAAACAAATAATTA 865  
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
Db 866 GAAGAAGCTAAAGAAATTTGGGATAAAGAGCTATCAGGCCAACATTTCTATTGTGGCC 925  
Qy 259 ValTyrPhePheMetAsnGlyTyrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278  
Db 926 GCCTTCTTATTGATCATGCATCATATGCTCTGGCTTTCTGGTATGGACCTCTCTGGTC 985  
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
Db 986 CTCTCCAGTGAA-----TATTCTATTGGACAAAGTACTCACTGTCTTCTTCTGATTATTA 1039  
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
Db 1040 ATTTGGGCTTTTGTAGTATTGGACGGCATCCCCAGCATTTGAAGCATTTGCAACGCAAGA 1099  
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
Db 1100 GGAGCAGCTTATGAATCTTCAAGATAATTGACAATAAACCAAGCATTTGACAGCTATTCCG 1159  
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358  
Db 1160 AAGAGTGGACATAAACCAGATATATTAAAGGAAATTTTGGAAATTCAAAATGTCTCACTTC 1219  
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
Db 1220 AGTTACCTTCTCGAAAAGAAAGATTAAAGATCTTAAAGGGTCTCAACCTGAAGTTACAGT 1279  
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
Db 1280 GGGCAGACAGTGGCGTGGTGGGAACAGTGGCTGGGGAAGACGACGACCGTGCAGCTG 1339  
Qy 399 LeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArg 418  
Db 1340 ATGACAGAGGCTTATGACCCACAGATGGCATGGTCTGTATTGATGGACAGGACATAGG 1399  
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlnProValLeu 438  
Db 1400 ACCATAATGTAAGGCATCTTCGGGAAATTTACTGGTGTGTGTAGTCAGGAGCTGTGTTG 1459  
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458



Db 1460 TTTGCCACCATGATGCTGAAACATTTCTGCTATGGCGCGAAATGTCACCATGGATGAG 1519  
QY MetGluArgAlaAlaargGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
Db 1520 ATTGAGAAAGCTGTTAAGGAGCCATGCTCTATGATTTATCATGAACATACCTAATAAA 1579  
QY PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgile 498  
Db 1580 TTTGACACTCTGTTGGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639  
QY AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
Db 1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTTCTTCTGGATGAGGCAACGTCA 1699  
QY AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538  
Db 1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGTGGCCCTGGATAAGGCCAGAAAGGC 1759  
QY ArgThrThrIleValValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558  
Db 1760 CGGACTACCATTTGATAGCTCATCGTTTGTCTACAGTTCCGTAATCCGATGCTATTGCT 1819  
QY ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlnLeuMetAlaLysArg 578  
Db 1820 GGTTTTGTATGGAGTCACTGTGGAGAAAGGAAATCATGATCAACTCATGAAGAGAGAAG 1879  
QY GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598  
Db 1880 GGCATTACTTCAAACTTGTCACAAATGCAG---ACAAGAGAAATGAAATGAGTTAGTAA 1936  
QY SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
Db 1937 AATGCCACTCGTGAATCCAAAGTGAAGAGTGCCTTGGAAATG---TCTCCAAAGAT 1993  
QY IleLysSerAspPheIle----- 624  
Db 1994 TCAGGGTCCAGTTTAAATAAAGAGATCAACTCGCAGGAGTATACATGCACCACCAAGGC 2053  
QY 625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643  
Db 2054 CAAGACAGAAAGCTTGGTACAAAGAGGACTTGATGAGAAATGTACTCTCAGTTTCTTTC 2113  
QY LeuLysIleLeuLysLeuAsnLysProGluTrpPheValValLeuGlyThrLeuAla 663  
Db 2114 TGGAGGATTCTGAAGCTGAACCTCAACTGAATGACCTTATTTTGTGGTGTATATTTGT 2173  
QY SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683  
Db 2174 GCTATTATAACCGAGGCCCTGCAACCAAGCATTTTCAATAATATATTTCAAGGATATAGGG 2233  
QY MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702  
Db 2234 ATCTTTACCGAGATGAGGATCTCTGAAACAAACGACAGAAATAGTAACATGTTTCTGTA 2293  
QY IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722  
Db 2294 TTGTTTCTAGTCTCTGGAATTTTCTTTTATATACATTTTCTCCAGGGCTTCACATTT 2353  
QY GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisIleuAlaPheLysAlaMetLeu 742  
Db 2354 GGCAAGCTGGGAGATCCTCACTAAGCGGCTTCGATACATATGTTTTCAGATCCATGCTG 2413  
QY TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrIle 762  
Db 2414 AGACAGATGTCAGCTGGTTGATGACCCCTAAACCAACCATCGGAGCATTTGACAAACAGG 2473  
QY LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgileGlyValLeuThr 782  
Db 2474 CTTGGCCNATGATCGCGCTCAAGTTAAAGGGGCTATAGTTCCAGGCTTGTGTCATTTACC 2533  
QY GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802

Db 2534 CAGATATATAGAAAATCTTGGGACAGGCATTTATATATCTTAACTATCTATGTTGGCAATTA 2593  
QY ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822  
Db 2594 ACATTTTACTCTTACCAATTTGACCATCATTTGCAATGACAGGAGTTGTTGAAATGAAA 2653  
QY AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842  
Db 2654 ATGTTGCTCTGGCAAGACACTGAAAGATAAGAAAGACTAGAAAGGAGCTGGGAAGATTGCT 2713  
QY ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862  
Db 2714 ACAGAACCATCGAAACCTCCGAACTGTTGTTCTTTGACTCGGAGGACAGAGTTTGA 2773  
QY GlnMetTyrGluGluMetLeuGlnHisArgAsnThrSerLysLysAlaGlnIle 882  
Db 2774 TACATGTATGCACAGAGTTTGAAGTACCATACAGAACTCTTTGAGGAAAGCACACATC 2833  
QY IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902  
Db 2834 TTGCGGGTCTCATTTTCTATCACCAGGCAATGATATATTTTCTTATGCTGCTGCTTTC 2893  
QY ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922  
Db 2894 CGGTTTGGTCCCTACTTGTGGCAATGATTCATGAACTTTTCAGGATGTTCTTTTGGTA 2953  
QY PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942  
Db 2954 TTCTCAGCTATTGCTTTGTCATGGCAGTGGCAGTGGGCGAGTCACTGTTCAITTTGCTCTGAC 3013  
QY TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962  
Db 3014 TATGCCAAAGCCAAAGTATCAGCAGGCCACGTCATCATCATCATTTGAAAAAAGCCCTCTG 3073  
QY IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyLeuAsnLeuGluPhe 982  
Db 3074 ATTGACAGCTACAGCCCTCAGCGCCTCAAGCCAAATACGTTGGAAGAAATGTGACATTT 3133  
QY ArgGluValSerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002  
Db 3134 AATGAGGTCTGTTTCAACTATCCACTCGACCAGACATCCCGTGTCTCCAGGGGCTGAGC 3193  
QY LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022  
Db 3194 CTGAGGTGAAGAGGCGCAGACGCTGGCCCTCGTAGTAGCAGTGGCTGTGGGAAGAGC 3253  
QY ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042  
Db 3254 ACAGTTGTTGAGTCTCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGTAAATTGAT 3313  
QY GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062  
Db 3314 GGCAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCACTCGTGTCT 3373  
QY GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082  
Db 3374 CAGGAGCCCATCTGTTTGACTGTCAGCATTTGCCGAAACATTTGCCCTATGAGAGCAACAGC 3433  
QY ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102  
Db 3434 CGGGTCTGATTCATCAAGAGATTTATGAGGAGCCCAAGAGGCGCAATACACCATTC 3493  
QY IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122  
Db 3494 ATCGAGACACTCTCTGAGAAATACACACAGAGTAGGAGACAAAGAGNACCCAGCTCTCT 3553  
QY GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142  
Db 3554 GGTGGCCAGAAACAGCGCATTTGCCATAGCTCGCGCTCTTGTGTAGACAGCTCATATTTTG 3613  
QY LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162  
Db 3614 CTTTGGATGAAGTACATCAGCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAGCC 3673

Qy 1163 LeuAspLysAlaArgThrGlyThrCysLeuValValThrHisArgLeuSerAlaIle 1182  
 Db 3674 CTGGACAAAGCCAGAGAGCGCACCTGTCATTTGGATCGCCACCGCTTTCACCATC 3733  
 Qy 1183 GlnAsnAlaLeuLeuValValLeuHisAsnGlyLysLeuGlnGlyThrHis 1202  
 Db 3734 CAGAATGCAGATTTTAAATAGTGTCTTTCAGAAATGCCAAAGTCAAGAGGATGCGCACAT 3793  
 Qy 1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220  
 Db 3794 CAACAGCTGTGCGCCAGAAAGGATCTATTTTCCATGTGTCAGTCCAGGCT 3847

RESULT 13  
 AX105080 4279 bp DNA linear PAT 30-APR-2001  
 LOCUS  
 DEFINITION Sequence 24 from Patent WO0123540.  
 ACCESSION AX105080  
 VERSION AX105080.1 GI:13921230  
 KEYWORDS  
 SOURCE dog.  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE 1 (bases 1 to 4279)  
 AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.  
 TITLE P-glycoproteins and uses thereof  
 JOURNAL Patent: WO 0123540-A 24 05-APR-2001;  
 GENTEST CORPORATION (US)  
 FEATURES  
 source 1..4279  
 /organism="Canis familiaris"  
 /db\_xref="taxon:9615"  
 CDS  
 17..3862  
 /notes="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC37722.1"  
 /db\_xref="GI:13921231"  
 /translations="MPEGRGKSAENFWKMKKKKKKKKKPTVSTFAMFRYSN  
 WDLRLVLMVGTMAIITHGALPLMLVFNMDTSFANAGISRNKIPFVILINESITNNT  
 QHFNLHEEMTTIAYYIGIGAGVLAAYIQVFWCLAGROILKIRKOFTHAMRQ  
 EIGWFDVHDVGLNLTLDVSKINEGIDKIGMFQSIATFTFTGTFVGFTHAKLTL  
 VILAISPLVLSAAIKWILSSFTDKELLAYAKAGAVEVLAARITVAFGQKKEL  
 ERYKNLEAKGIGIKKATANISICAAELLIVASVALAFWTGTSVLSSYSIGOV  
 TVPFSVLGAFSGPSIEAPANARGAAYEIFKILDNKPSIDSKSGHPRDNKIG  
 NLEFNHFSYPSRKEKILKLNKQSGQTVLVNCGSKSTTVQLMQLYDPTD  
 GWICDQDITINVRHREITGVWSQEPVLFTATTIAENIRYGRNVNWDLEAKVE  
 SEAVOVOLDKARKGRTTVIAHRLSTVRNADVIAGDDGVIVKGNHDELMEKGIY  
 ANAYDFIMKLPNKFDLVGERGALSGQKQRIARALVRNPKILLDEATSLDTE  
 FKLVMTQRTNETLEENATCESKESDALEMSPKDSGLIKRRSTRRSIHAPQODR  
 KLGTEDLNENVPVFWRLKLNSTEWYFVVGIFCALINGLQAPFSLIIFSRIGI  
 FRDEDPETKRONNSMFSVLFLVLHISFITFFLOQTFGKAGEILTKRLRYVFRSM  
 LRQVSWDFDPKNTGALTRLANDAAVKGAISRLAVITQNIANLGTGIIISLYG  
 WOLTELLVAIPLIALGVVEMKLSQALKKKELEGAGKATIEFNFRVWLSLR  
 EOKFELMYAQSLOPVVRNSLRKAHIFGVFSITQAMVFSYAGCFRFGAVLVANEMN  
 FODVLVFNINVEGANVQVVSFADYAKAKVSAHVIMIEKSLIDSYPHGLKP  
 LDLEGNVTNEVFNPTPTDIPVLQGLSEKQTLALVSGGCKSTVQVLLERF  
 YDPLAGSLVDIGKHLNVQMLRAHLGIVSQEPKILFDCSIAENATIGNSVWSHEE  
 IMQAKENAIHHFETLPKYNTRVGDKTOLSGGOKRIARALVRPHQLLLDEA  
 TSLADTESKRVQALDKREGRTICIVIAHRLSTIQNALDILVVFQNGKVKHGTQOL  
 LAQKGIYFMSVSVQAGAKR"  
 BASE COUNT 1296 a 833 c 1008 g 1142 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4,45e-227 Length: 4279  
 Score: 3530.50 Matches: 682  
 Percent Similarity: 73.13% Conservative: 238  
 Best Local Similarity: 54.21% Mismatches: 269  
 Query Match: 57.32% Indels: 69  
 DB: 6 Gaps: 9

US-09-873-409-5 (1-1222) x AX105080 (1-4279)  
 Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 167 ATGTTGGTGGGACATGCTGCTCATCATCCATGGAGCTGCATCCCTCTCATGATGCTG 226  
 Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34  
 Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAAACCTTT 286  
 Qy 35 ValGlnThrAsnThrTyrSerPhePheArg 44  
 Db 287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACTTTCATCAACCATCTGGAG 346  
 Qy 45 LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
 Db 347 GAGGAATGACACGATGCTTATTATACAGTGGATGCTGCTGGCGTGGTGGTGGCT 406  
 Qy 60 GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
 Db 407 GCTTACATCCAGGTTTCATCTGCTGCTGGCAGCAGGAGACAGATACATCAAAATAGA 466  
 Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
 Db 467 AAACAATTTTTCATGCTATCATGCGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 526  
 Qy 100 GlyGluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAsp 118  
 Db 527 GGGAGCTTAACACCGGCTCACAGAGATGCTCCAAATCAATGAAGAAATGGCGAC 586  
 Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
 Db 587 AAAATTTGGAATGTTCTTCAATCAATAGCAACATTTTCCACCGGTTTATAGTGGGTTT 646  
 Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
 Db 647 ACACGTGGTGGAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGAGTTTCA 706  
 Qy 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
 Db 707 GCGCCATCTGGCAAGATATCTTCTTACTTGTATAAGAACTCTTGGCCTATGCA 766  
 Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
 Db 767 AAGCTGGAGCAGTACTGAAGAAGCTTAGCAGCAATCAGAACTGCTGATGCTTGA 826  
 Qy 199 AlaGlnGlyLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe 218  
 Db 827 GGACAAAGAAAGAACTTGAA 847  
 Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeuPheValArgTyrThrGlnAsnLeu 238  
 Db 848 AGGTACACAAAAATTTA 865  
 Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
 Db 866 GAAGAAGCTAAGGAATGGGATAAAGAAAGTATCAGCGCAACATTTCTATGTTGCC 925  
 Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278  
 Db 926 GCTTCTTATGATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 985  
 Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
 Db 986 CTCTCCAGTGAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1039  
 Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
 Db 1040 ATTGGGGCTTTTATGATTTGGACAGGATCCCAAGCATTTGAGCATTTGCAACGCAAGA 1099  
 Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysGlyProSerIleAspAsnPheSer 338  
 Db 1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGCAATAAATGAAGCATTTGACAGCTATTGCG 1159



```

Db 3314 GCAAAGAGATAAGACCTGAATGTCCAGTGTCCGAGCACACCTGGCGCATCGTGTCT 3373
Qy 1063 GlnGluProValLeuPheAenCysSerIleAlaGluAenIleAlaTyrGlyAspAenSer 1082
Db 3374 CAGGAGCCATCTCTGTTGACTCAGCATTCGCCGAGAAATTCCTTATGAGACAACAGC 3433
Qy 1083 ArgValValProLeuAspGluIleLysGluAlaAlaAenAlaAlaAenIleHisSerPhe 1102
Db 3434 CGGGTGTATCATGAGAGATTTATGAGGAGCCAGGAGGCCAATACACCACTTC 3493
Qy 1103 IleGluGlyLeuProGluLysTyrAenThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db 3494 ATCCAGACACTCCCTCAGAAATACACACAGAGTAGGAGACAAAGAGAACCCAGCTCTCT 3553
Qy 1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaAenAlaLeuGlnLysProLysIleLeu 1142
Db 3554 GTGGCCAGAAACAGCGCATGTCATAGCTCGCGCTCTTGTATGACAGCTCTATATTGG 3613
Qy 1143 LeuLeuAspGluAlaThrSerAlaLeuAspAenAspSerGluLysValValGlnHisAla 1162
Db 3614 CTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAGGTTGTCCAAGAGCC 3673
Qy 1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db 3674 CTGGACAAGCCAGAGAGCCGACCTGATGTCATGATCGCCACCGCTTGTCCACCATC 3733
Qy 1183 GlnAsnAlaAspIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db 3734 CAGAATGCAGATTAATAGTGGTGTTCAGAAATGGCAAGTCAAGGAGCATGCGACAT 3793
Qy 1203 GlnGluLeuLeuArgAenArgAspIleTyrPheLysLeuValAenAlaGlnSer 1220
Db 3794 CAACAGCTGTGCCCAAGAGCACTATTTTCATGTGTGATGTCAGGTCTCCAGGCT 3847

RESULT 14
AX391099
LOCUS 4646 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0210205.
ACCESSION AX391099
VERSION AX391099.1 GI:19584228
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Ruth,A. and Roninson,I.
Mutations of the mdrl p-glycoprotein that improve its ability to
confer resistance to chemotherapeutic drugs
Patent: WO 0210205-A 1 07-FEB-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
source
    1. 4646
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    425_4267
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD28599.1"
    /db_xref="GI:19584229"
    /translation="MDLEGRNGAKKKKFFLNKSEKDKKKKPTVSVFNMFRYSN
    WLDKLVVVGTLAAIITHAGLPLMLVFGEMTDIIPANAGNLELMSNIINRSDINDTG
    FFMNLEDMTRVAYYSGGTGVAGVVAAYIQVSPWCLAAGRQHKIKQFPHALMRQEI
    GMPDHDVGEINRLTDDVSKNEGIDGKIMFPFQSMATFFGTFVGRFGWKLILVI
    LAISPVGLSSAAWMAKILSFTDKELLYAKAGAAVEAIAITVFAGGKKLEL
    YNNLEAKRIGIKKAITANISIGAAFLIYALAFWYGTTLVLSGSYSGQVLT
    FRSVLIGATVGSPIISAFANARGAAVEIFKIDNKPISIDSKSGHKPNKGNL
    EFNHVSFSPKREVKILGKLNKVQSGOTVALVGNSSGCKSTTVOLMORLYDPTEGM
    VSDVDGNTINVRFLREIIGVVSQSPVLPATIAENIRYGRNVTMDSEIAKVEAN
    AYDFIMKLPHKFTLVGERGAQUSGQKQRIARALVRNPKILLDLDEATSDTSE
    AVQVQALDKARQGTITVIAHRLSTVRNADVAGFDGQVIVKGNHLELMKRGYFK
    LVMTQTAGNEVELENAADESKSEIDALEMSSNDSRSSLIIRKSTRSRVRSQAQRKL

```

BASE COUNT 1372 a 892 c 1130 g 1252 t

ORIGIN

Alignment Scores:

Pred. No.: 4,95e-227 Length: 4646  
 Score: 3530.50 Matches: 682  
 Percent Similarity: 74.00% Conservative: 246  
 Best Local Similarity: 54.39% Mismatches: 263  
 Query Match: 57.32% Indels: 63  
 DB: 6 Gaps: 10

US-09-873-409-5 (1-1222) x AX391099 (1-4646)

Qy 1 MetIleLeuGlyIleLeuLysSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 575 ATGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACATCTCTCATGATGCTG 634  
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 635 GTGTTGGAGAAATGACAGATATCTTTGCAATGCGAGAAATTTAGAGATCTGATGTCA 694  
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 695 ACATCATCTAATAGAGTATCATGATACAGGGTCTTCTCATGATCTGGAGGAGAC 754  
 Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 755 ATGACCAGGTATGCTTATTATACAGTGGAAATTTGGTCTGGGGTGTCTGTTGCTTAC 814  
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 815 ATTCAGGTTTCATTTTGGTGGCTGGCAGTGGAGACAAATACACAAATTTAGAAACAG 874  
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
 Db 875 TTTTTCATGCTATAATCGACAGAGATAGGTGCTGTTGATGTCACCATGTTGGGAG 934  
 Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 935 CTTAACACCCGACTTACAGATGATGCTCTCTAAGATTAATGAAGGAATTTGGTGACAAA 994  
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 995 GGAATGTTCTTTCAGTCAATGGCAATTTTTCATGCGGTTTATAGTAGATTACACGT 1054  
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 1055 GGTTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTTCAGCTGCT 1114  
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 1115 GTCTGGGCAAGATACATCTCTTCAATTACTGTATAAGAAATCTCTTAGCGTATGCAAA 1174  
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACCTGTGATTCATTTGAGAGCAA 1234  
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
 Db 1235 AAGAAAGAACTTGAA----- 1249  
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240

STKEALDESIPPVSWFRIMKLNLTWEPYFVVGVFCALINGLQPAFAIIFSKIIQVFT  
 RIDDPETKRONSNLFSLLFLALGIIISFTIFLQGTFTGKAGEIILTKLRYWVFRSMLR  
 QDVSWDDPKNTGALTFLANDAAQVKGATGSRNAVITQNIANLGTGIIISFTYGWQ  
 LTLALLAIVPIAIAGVVMKMLGQKDKKELEGAGKATEAIENFRNVTVSLTQEO  
 KFEHWAQSLQVPYRNSLRKAHIGITFTSQMMYFSYACGFRGAYLVAKHLSMFE  
 DVLVFSAVVFGANAVGVSSFADYAKAKISAAHILMIJEKTLIDISYSTGELMPNT  
 LEGNVTGEEVFNYPTRPDIPVLOGLSLEVKKGTLLALVSGSGCKSTVQVLLERFYD  
 PLAGKLLLDGKEIKRLNVQMLRAHLGVISQBPILFDSCISAENIAYGDNRSRVSEIV  
 RAAKEANIHFPIESLPNKYSTKGDKGTQLSGGQKRIAIARALVRQPHILLDEATS  
 ALDTSEKVOEALDKAREGRTCVIAHRLSTIQNADLIVVFQNGRVRKEHGTQOLLA  
 QKGIYFSMVSVQAGTKRQ"



Qy	947	LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg	966
Db	3428	AAATATCATGACGCCACATCATCATGATCATGTTGAAAACCCCTTGATTGACAGCTAC	3487
Qy	967	SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer	986
Db	3488	AGCAGGAAGCCCTAATGCCCAACATTCGGAAGAAATGTCACATTTGGTGAAGTTGTA	3547
Qy	987	PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu	1006
Db	3548	TTCAACTATCCACCCAGCCGACATCCCGAGTCTTCAGGAGATGACCCCTGGAGGTGAAG	3607
Qy	1007	ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln	1026
Db	3608	AAGGCCACAGCGTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAG	3667
Qy	1027	LeuLeuGlnArgLeuTyrAspProValGlnGlyClnValLeuPheAspGlyValAspAla	1046
Db	3668	CTCTGAGCGGTTCTACGACCCCTTGGCAGGAAAGTGTCTGTGATGGCAAGAAATA	3727
Qy	1047	LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal	1066
Db	3728	AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATC	3787
Qy	1067	LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro	1086
Db	3788	CTGTTTACGTGCAGCATGCTGAGAACATTCCTATGAGACACACACCGGGTGTGTC	3847
Qy	1087	LeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu	1106
Db	3848	CAGGACAGATCGTGAGCGCAGCAAGAGGAGCCCAACATACATGCTTCATCGAGTCACTG	3907
Qy	1107	ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys	1126
Db	3908	CCTAATAATATACATTAAGTAGGAGACAAAGAACTCAGCTCTCTGTGGGCCAGAAA	3967
Qy	1127	GlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGlu	1146
Db	3968	CAAGCATGCCATAGCTCGTGCCTTGTAGACAGCCTCATATTTTGTCTTGGATGAA	4027
Qy	1147	AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla	1166
Db	4028	GCCACGTCAGCTCTGGATACAGAAAGTGAAGAGTGTCTCCAAAGAGCCCTGGACAAAGCC	4087
Qy	1167	ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp	1186
Db	4088	AGAGAAGCCGACCTGATGTGATGTGCTCACCCTGTCTCCACCATCCAGAAATGCAGAC	4147
Qy	1187	LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu	1206
Db	4148	TTATATAGTGGTTCAGATGGCAGAGTCAAGGAGCATGCGACGCATCAGCAGCTGCTG	4207
Qy	1207	ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer	1220
Db	4208	GCACAGAAAGGCATCTATTTTCAATGGTCAAGTCCAGGCT	4249
RESULT	15		
AF269224		4045 bp	DNA linear SYN 04-JUL-2000
LOCUS			
DEFINITION			Synthetic construct Canis familiaris his-tagged-multidrug
			resistance glycoprotein gene, complete cds.
ACCESSION			AF269224
VERSION			AF269224.1 GI:8926216
KEYWORDS			synthetic construct.
SOURCE			synthetic construct
ORGANISM			artificial sequences.
REFERENCE			1 (bases 1 to 4045)
AUTHORS			Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE			Analysis of dog MDR1 p-glycoprotein
JOURNAL			Unpublished
REFERENCE			2 (bases 1 to 4045)
AUTHORS			Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE			Direct Submission
JOURNAL			Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chamin de
			Tournefeuille, BP3, Toulouse 31311, France
FEATURES			Location/Qualifiers
source			1..4045
			/organism="synthetic construct"
			/db_xref="taxon:32630"
			/note="original mRNA isolated from Beagle dog colon"
			101..3964
			/codon_start=1
			/transl_table=11
			/product="his-tagged-multidrug resistance glycoprotein
			MDR1"
			/protein_id="AAF81747.1"
			/db_xref="GI:8926217"
			/translation="MHHHHHDPGGGRKGSAGKKNFKWKGKSKKNEKKEKKPTVSTFA
			MPFRYNWLDRLYMLVGTWAAIIHGALPLMLVFCNMTDSPANAGISRNKTPPVLIN
			SITNTQHFINHLEEMTYAYYSIGIGAGVLVAAYIQVSPHCLAGRQILKIRQFF
			HAIIMRQIGWDFVDHDLNRLTDDVSKINIGIDKIGMFPQSITATFTTFFIGVTR
			GOKLRLYLAISPVILGSAAIWAKILSSFTDKELLYAKAGAVABEVLAIRTVAFG
			GOKLELYNKLEAKRIGIKKAITANISIGAAFLIYASVALAFWTGTSVLSEY
			STGOVLTVFFSVLIGAFSIGOAPSPSIEAFANARGAAVEIFKIDNKPISDSYKSHK
			PDNIKGNLEFNKLVHESYSPRKEVKILGILNKVQSGQTVVALVNSGCGKSTVOLMOR
			LIDPDMGWDGQDIRIINVRHLREITGVVSQEPVLEATTIAEMIRYGRNVTWDEI
			EKAKEANAYDFIMKLPNKFDPLVGERGAQSGGQKQRIATIALVRNPKILLDEAT
			SALDTSEAVVQVADLKARQRTTIVIAHRLSTVRNADVIAGDFDGVIVKGNHDELM
			KEGIVFKLVMTQNGEILEENATGESKSDALEMSPKDSGSSILKRRSRSTHA
			POGQRKLGTEDLNENVPVFWIRILKINSEWPFYVVGIFCAINGLQPAFSIIF
			SRIGIGTRDEDPETKRONSNMFSVLVLVGLIISFITPFGQTFGKAGEIILKRLY
			MYFRMLKQDVSDFDDPNTTALTTRLANDAAQVKGALIGSLAVITQNIANLGTGII
			ISLIYQWQLTLLLAIPILAIAGVEMKUSQALKKKLEIGELGKQRIATEALENPT
			VVSLREQFEYMAQSLQVPMYNSRLKAHIFGVSFITQAMMYFACGFRGAYLV
			ANEFNFODLIVFSAIVFGMAVQSVSFADYAKAKVSAAHVIMIIEKSLIDSYS
			PHLKPNTLEGVNTNEVFNTPDPVLQGLSLEVKKGOTLALVGSGCCSTVW
			QLLERPYDPLAGSVLIDGKELKHLNVOMLRAHLGIVSOEPIFDCSIAGNIAYGDSNR
			VVSHETIQAARKEAHNHFIEITPEKYNTRVGDKGTQLSGGKQRIARALVRQPHI
			LLLDTSALDTESEKVOEALDKAREGRTICVIAHRLSTIQNADLIIVFQNGKVAEH
			GTHQQLLAQKGIYFMSIVQAQK"
BASE COUNT		1171 a	850 c 986 g 1038 t
ORIGIN			
Alignment Scores:			
Pred. No.:	5,64e-227	Length:	4045
Score:	3528.50	Matches:	681
Percent Similarity:	73.13%	Conservative:	239
Best Local Similarity:	54.13%	Mismatches:	269
Query Match:	57.29%	Indels:	69
DB:	12	Gaps:	9
US-09-873-409-5 (1-1222) x AF269224 (1-4045)			
Qy	1	MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMatProLeu	20
Db	269	ATGTGTGGTGGGACAATGGTGCATCATCCATGGAGCTGCATCCCTCTCATGATCTG	328
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu	34
Db	329	GTTTTGGAAACATGACAGATAGCTTTGCAATGTCAGGAATTTCAAGAAACAAACTTTT	388
Qy	35	-----ValGlnThrAsnThrTyrSerPhePheArg-----	44
Db	389	CCAGTTATAATTAATGAAGTATTATACGAACATACACACATTTTCATCAACCATCTCGAG	448
Qy	45	-----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe	59
Db	449	GAGGAATGACACGATGCTATTATACGTGGGATCGGTGCTGGCTGCTGGTCTGCTG	508
Qy	60	GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg	79
Db	509	GCTTACATCCAGGTTTCATTCTGTGCTGCTGCGAGCAAGACAGATACATCAAAATAGA	568
Qy	80	LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle	99



Db 569 AAACAATTTTTTCATCTATCATCGACAGGAGATTGGCTGTTTCACGTGCATGACGTT 628  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 629 GGGGAGCTTAAACACCGGCTCACAGCAGTGTCTCAAAATCAATGAAGAAATGGCGAC 688  
Qy 119 LysIleAlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 689 AAAATTGGAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 748  
Qy 139 ValLysGlyTriPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 749 ACACGTGGTGGAGGTAAACCTTGTGATTTTGGCCATCAGCCCTGTTTGGCATTTCA 808  
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 809 GCGCCCATCTGGCAAGATATCTATCTTACTATGATAAAGAACTCTTGGCTATGCA 868  
Qy 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCTTTGGA 928  
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218  
Db 929 GCACAAAGAAAGAACTTGAA----- 949  
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsnLeu 238  
Db 950 -----AGGTACAAACAAAATTTA 967  
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
Db 968 GAAGAAGCTAAAGAAATTCGGATAAAGAAAGCTATCACGCCCAACATTTCTATTGGTGC 1027  
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278  
Db 1028 GCCTTCTTATGATCATCATCATATGCTCTGGCTTCTGGTATGGGACCTCTCTGGTC 1087  
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
Db 1088 CTCCTCAGTGAA-----TATTATTTGGCAAGTACTACTCTCTCTTTCTGTTATTA 1141  
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
Db 1142 ARTGGGCGCTTTAGTATTGACAGGCTATCCCAAGCATTTGAACATTTGCAACGCAAGA 1201  
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
Db 1202 GGAGCAGCTATGAATCTTCAAGATAATTGACAATAAACAACCAAGCATTTGACAGCTATTCG 1261  
Qy 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358  
Db 1262 AAGAGTGGACATAAACACAGATAATATAAGGGAAATTTGGAAATTCAAAATGTTCACTTC 1321  
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
Db 1322 AGTTACCTTCTCGAAAAGAAAGTTAGATCTTAAAGGTTCTCAACCTGAAGTTTCAGAGT 1381  
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
Db 1382 GGGCAGACAGTGGCGCTGTTGGGAACAGTGGCTCGGGAAGACGACCGCTGCAGCTG 1441  
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
Db 1442 ATCCAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTATGACAGGACATAGG 1501  
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
Db 1502 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGGTGTGTGAGTCAGGAGCCTGTGTTG 1561  
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458  
Db 1561 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458

Db 1562 TTTGCCACCAACGATAGCTGAAACCAATTGCTATGCGCGCGAAAAATGTCTACCATGGATGAG 1621  
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
Db 1622 ATTGAGAAAGCTGTTAAGGAAGCCCAATGCTATGATTTTATCATGAACACTACCTAATAAA 1681  
Qy 479 PheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
Db 1682 TTTGACACTCTGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAAATC 1741  
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
Db 1742 GCATTTGCTGGGCCCTGTTGCGAACCCCAAGATTTCTTCTGCTGATGGCAAGCTCA 1801  
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538  
Db 1802 GCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAAGTGGCCCTGGATAAGGCCAGAAAGGC 1861  
Qy 539 ArgThrThrIleValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558  
Db 1862 CGGACTACCATTTGTGATAGCTCATGCTTTGCTACAGTTTCTGAATGCCGATGTCATTGCT 1921  
Qy 559 ThrLeuLysAspGlyMetLeuAlaGlyLysGlyAlaHisAlaGluLeuMetAlaLysArg 578  
Db 1922 GGTTTTGATGAGTGGATCATTTGAGAGAAAGAAATCATGTAACACTCATGAAGAGAAG 1981  
Qy 579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLeuLysAlaAspGluGlnMetGlu 598  
Db 1982 GGCATTACTTCAAACTTGTCAATGTCAG---ACAAGAGGAAATGAAATTTAGTTAGAA 2038  
Qy 599 SerMetTyrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618  
Db 2039 AATGCCACTGCTGCAATCCAAAAGTGAAGTGAATGATGCTTGGAAATG---TCTCAAAAGAT 2095  
Qy 619 IleLysSerAspPheIle----- 624  
Db 2096 TCAGGCTCCAGTTTATAAAAAAGAAAGATCACTCGCAGGAGTATACATGCACCAAGGC 2155  
Qy 625 ---AspLysAlaGluLeuSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643  
Db 2156 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACTCCAGTTTCTCTTC 2215  
Qy 644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663  
Db 2216 TGGAGGATTTCTGAAGCTGAACCTCAACTGAATGCCCTTATTTTGTGGTGTATATTTGT 2275  
Qy 664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683  
Db 2276 GCTATTATAACCGAGGCTCGCAACCAAGCATTTTCAATAATATTTTCAAGGATATATAGG 2335  
Qy 684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702  
Db 2336 ATCTTTACCGAGATGAGGATCTCTGAAAACAAACGACAGATAGTAAACATGTTTCTGTA 2395  
Qy 703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722  
Db 2396 TTGTTTCTAGTCTTGGAAATTTATTTTATTAATTTTCTTCCAGGCTTTCACATTT 2455  
Qy 723 GlyArgAlaGlyLeuIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742  
Db 2456 GGCAAAGCTGGGAGATCTCTCACTAAGCGGCTTCGATATGTTTTCAGATCCATGCTG 2515  
Qy 743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762  
Db 2516 AGCAGGATGTAGCTGGTTTGTATGATCCCTAAAAACACCACTGGAGCATTGACACAGG 2575  
Qy 763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782  
Db 2576 CTTGCCAATGATCGGCTCAAGTTAAAGGGCTATAGTTTCCAGGCTTGTGTCATTACC 2635  
Qy 783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802  
Db 2636 CAGAAATATAGCAAACTCTGGGACAGGCAATTTATATATCTTATCTTATGTTGTCGAATTA 2695





GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 434.971 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012

Sequence: 1 MILGILASLVNAGCLPLMPL.....QELLNRDIYKLVNAOSVQ 1195

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spo3/US09873409/runat\_27032003\_115424\_19290/app\_query.fasta\_1.7544  
-DB=Published Applications NA -Qfmt=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=blom62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THL MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09873409 @CCN 1.1 960 @runat\_27032003\_115424\_19290  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	DB ID	Description
1	6012	100.0	3621	US-09-873-409-14
2	5988.5	99.6	3702	US-09-873-409-13
3	5159.5	85.8	3177	US-09-873-409-12
4	4561.5	75.9	2856	US-09-873-409-10

5	3554	59.1	3860	10	US-09-866-866A-1	Sequence 1, Appli
6	3549	59.0	4317	9	US-10-044-671-1	Sequence 1, Appli
7	3545	59.0	3860	10	US-09-866-866A-3	Sequence 3, Appli
8	3545	59.0	8630	10	US-09-306-417-1	Sequence 1, Appli
9	3545	59.0	8630	10	US-09-306-417-2	Sequence 2, Appli
10	3540.5	58.9	4425	10	US-09-769-097-3	Sequence 3, Appli
11	3538.5	58.9	4369	10	US-09-769-097-1	Sequence 1, Appli
12	3537.5	58.8	4643	9	US-10-072-621-2	Sequence 2, Appli
13	3526.5	58.7	4189	10	US-09-866-866A-7	Sequence 5, Appli
14	3522.5	58.6	4788	10	US-09-866-866A-7	Sequence 7, Appli
15	3504.5	58.3	3924	10	US-09-880-107-2299	Sequence 2399, Ap
16	3476.5	57.8	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
17	3456	57.5	2066	10	US-09-873-409-9	Sequence 9, Appli
18	3428.5	57.0	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	2835	47.2	5010	10	US-09-917-800A-483	Sequence 483, App
20	2552	42.4	1940	10	US-09-873-409-16	Sequence 16, Appl
21	2528.5	42.1	2021	10	US-09-873-409-15	Sequence 15, Appl
22	2380.5	39.6	4175	10	US-09-749-340-3	Sequence 3, Appli
23	2351.5	39.1	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
24	2282.5	38.0	4002	9	US-09-758-828-1	Sequence 1, Appli
25	2282.5	38.0	4002	9	US-09-758-828-3	Sequence 3, Appli
26	2153	35.8	3861	9	US-09-938-842A-263	Sequence 263, App
27	2046	34.0	3512	10	US-09-749-340-7	Sequence 7, Appli
28	2010.5	33.4	4653	9	US-10-101-388-2	Sequence 2, Appli
29	1659.5	27.6	3792	9	US-09-882-694-10	Sequence 10, Appl
30	1617	26.9	2698	10	US-09-749-340-5	Sequence 5, Appli
31	1550	25.8	8036	9	US-10-101-388-1	Sequence 1, Appli
32	1538.5	25.6	2681	10	US-09-749-340-8	Sequence 8, Appli
33	1525	25.4	3999	9	US-09-882-694-9	Sequence 9, Appli
34	1102.5	18.3	1810	10	US-09-749-340-4	Sequence 4, Appli
35	1101.5	18.3	1175	10	US-09-873-409-11	Sequence 11, Appli
36	1084	18.0	7296	10	US-09-070-927A-59	Sequence 59, Appl
37	1031.5	17.2	3741	9	US-09-738-626-1251	Sequence 1251, Ap
38	1021	17.0	9047	10	US-09-070-927A-12	Sequence 12, Appl
39	1005	16.7	12438	10	US-09-070-927A-173	Sequence 173, App
40	941	15.7	64681	10	US-09-790-988-1	Sequence 1, Appli
41	902.5	15.0	3601	10	US-09-070-927A-255	Sequence 255, App
42	901.5	15.0	6415	10	US-09-070-927A-275	Sequence 275, App
43	781	13.0	2298	9	US-10-156-239-18	Sequence 18, Appl
44	781	13.0	2298	10	US-09-795-693-18	Sequence 18, Appl
45	781	13.0	3408	9	US-10-156-239-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1  
US-09-873-409-14  
; Sequence 14, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 3621  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-14

Alignment Scores:  
Pred. No.: 0  
Score: 6012.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 10  
Length: 3621  
Matches: 1195  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-873-409-6 (1-1195) x US-09-873-409-14 (1-3621)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 34 ATGATCTGGGTATACCTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAAATGCCACTG 93  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 94 GTTTTAGGAGAAATAGTGATTAACCTTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153  
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
DB 154 TCCTTCACAGTTGACCCCTGATTATGTTGGATAGGTGTGCTGCCTTGATTTTGGT 213  
QY 61 TyIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleValArgIleArgLys 80  
DB 214 TACATACAGATTTCTTGTGGATTATTAACCTGCAGCAGACAGACCAGAGATTCGAAAA 273  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
DB 274 CAGTTTTTTTCATTCACTGGTGGCAGACATCGGCTGGTTTGATAGCTGTGACATCGT 333  
QY 101 GluLeuAsnThrArgMetThrAspIleAspIleSerAspGlyIleGlyAspLysIle 120  
DB 334 GAACCTTAACACTCGCATGCAGACATTCACAAAATCAAGTATGAGATTAAGATT 393  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 454 GGCTGGAAACTCACCCTAGTGACTCTATCCAGCTCTCTCTTATATGGCTTCAGCGGCA 513  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 514 GCATGTTCTTAGGATGGTCATCTCATTCACAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 574 GGGCTGTGGCAGNAGAGCTTGTGTCATCAATCCGAACAGCATAGCTTTTAGGGCCAG 633  
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
DB 634 GAGAAAGAACTTCAAGGTATACACAGAACTCTCAAGATGCAAAAGGATTTTGGCATAAAA 693  
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
DB 694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTACTTCTTATGAATGGAACCTAT 753  
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
DB 754 GGACTTGCTTTTGGTATGNAACCTCTTGATCTTAAATGGAGNACCTGGATATACCATC 813  
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
DB 814 GGGACTGTCTTGCTGTTCTTTAGTGTAAATCCATAGCAGTTATTGCAATGGAGCAGCA 873  
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
DB 874 GTCCCTCACCTTGAACCTTCGCAATAGCCCGAGAGCTGCCTTTCATATATTTCCAGGTT 933  
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
DB 934 ATTGATAAGAAACCCAGATAGATAACTTTTCCACAGCTGATATAAACTCGAATCCATA 993  
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
DB 994 GAAGGAACCTGTGGAATTTAAAAATGTTTCTTCAATTAATCCATCAAGACCATCTATCAAG 1053  
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
DB 1054 ATTTCTGAAGAGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTGGTCTC 1113

QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
DB 1114 AATGGCAGTGGGAAGAGTAGTACGGTAGTCCAGCTCTTCAGAGGTTATATGATCCGGATGAT 1173  
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
DB 1174 GGCTTTTATCATGTGTGATGAGATGACATCAGAGCTTTTAAATGTGCGCATTTATCGAGAC 1233  
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
DB 1234 CATATTGGAGTGTGTAGTCAAGACCTGTTTGTTCGGACCCACCATCATGTAACATATC 1293  
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
DB 1294 AAGTATGGACGACATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGGAAGCAAAAT 1353  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
DB 1354 GCGTATGATTTTATCATGTGAGTTTCCATAAATTTAAATACATTTGGTAGGGGAAAAAGGA 1413  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
DB 1414 GCTCAATGAGTGGAGGGCAGAAACAGAGATCGCAATTTGCTCGTGCCTTTAGTTCGAAC 1473  
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
DB 1474 CCCAAGATTCTGATTTTATGATGAGGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCT 1533  
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
DB 1534 GTTCAAGCTGCATCTGGAGAAAGCGAGCAAAAGGTCGAGCTACAAATCGTGTGTAGCACACCGA 1593  
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
DB 1594 CTTTCTACTATTTCGAAGTGCAGATTTGATTGTGACCCTTAAAGGATGGAATGCTGGCGGAG 1553  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
DB 1654 AAAGGAGCAGCATCTCAACTAATGCAAAACAGAGGCTCTATATTATTCACTTGTGATGTCA 1713  
QY 561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580  
DB 1714 CAGGATATTTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAGAAAG 1773  
QY 581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600  
DB 1774 ACCACTCACTTCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGCAAGGCT 1833  
QY 601 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 620  
DB 1834 GAGGAATCCACCCCACTAAAGAGATAAGTCTTCTCTGAAGTCTCTCTATTAATAATTTTA 1893  
QY 621 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640  
DB 1894 AAGTTAAACAAGCCTCAATGGCCTTTTGTGGTTCTGGGACATTTGGCTTCTGTTCTAAAT 1953  
QY 641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 660  
DB 1954 GGAACCTGTTTCATCCAGTATTTCCATCATCTTTTGCAAAAATTAATAACCATGTTTGGAAAT 2013  
QY 661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680  
DB 2014 AATGATATAACACCATTAAGCATGATGCAGAAATTTATTCATGATATTCGTCATTTCG 2073  
QY 681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700  
DB 2074 GGTGTTATTTGCTTTCTCAGTTATTTTCATGCAGGGATTTATTTACGCGCAGCAGCGGAA 2133  
QY 701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720  
DB 2134 ATTTTAACGATGAGATTAAGACACTTGGCCCTTCAAGCCCAATGTTATATCATCAGGATATTGCC 2193

QY 721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740  
DB 2194 TGGTTTGGATGAAAGGAAACACGACAGGAGGCTTGACAAACATATTAGCCATAGATATA 2253  
QY 741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760  
DB 2254 GCACAAATTCAGAGGACACAGGTTCCAGATTGGCGTCTTAACACAAATGCAACTAAC 2313  
QY 761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeu 780  
DB 2314 ATGGGACTTTCAGTTATCATTTCTTTATATATGGATGGAGATGACATTCCTGATTCTG 2373  
QY 781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800  
DB 2374 AGTATTGCTCCAGTACTTGGCGTGACAGGAATGATTGAAACCCGACGAATGACTGGATT 2433  
QY 801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyIleAlaThrGluAlaLeuGlu 820  
DB 2434 GCCAACAAAGATAAGCAAGAACTTAGAGCATGTCTGGAAAGATAGCAACTGAAGCTTTGGAG 2493  
QY 821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840  
DB 2494 AATATACGTACTATATGTGTATTAACAGGGAAAGAGCTTCAGGCANATGTATGAAGAG 2553  
QY 841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860  
DB 2554 ATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATATTGGAAGCTGTAT 2613  
QY 861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880  
DB 2614 GCATTGAGCATGCCCTTATATATTTTGGCTATGCGACAGGGTTTCGATTGGAGGCTAT 2673  
QY 881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900  
DB 2674 TTAATTCAGCTCGACGATGACCCAGAGGGCATGTTCTATGATTTTACTGCAATTGCA 2733  
QY 901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920  
DB 2734 TATGGAGCTATGGCCATCGGAAAGCGCTCGTTTGGCTCCTGAATATTCGAAAGCCAAA 2793  
QY 921 SerGlyAlaAlaHisIlePheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 940  
DB 2794 TCGGGGGCTCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGACAGCCGACGT 2853  
QY 941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960  
DB 2854 CAAGAAGGAAAGGAAAGCCAGACATGGAAGGGAATTTAGAGTTTCGAAAGTCTCTTTC 2913  
QY 961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980  
DB 2914 TTCTATCCATGTCGCCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGA 2973  
QY 981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000  
DB 2974 GGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACATCTCTGTTCAACTT 3033  
QY 1001 LeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLys 1020  
DB 3034 CTCAGAGACTTTATGACCCCGTGCAAGGACAAGTGTCTTTGATGGTGGGATGCAAAA 3093  
QY 1021 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040  
DB 3094 GAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCCCAAGAGCGCTGCTC 3153  
QY 1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060  
DB 3154 TTCAACTGCGACGATTTGTGAGAACATCGCTATGTTGACAAACAGCCGCTGTGGTGCATTA 3213  
QY 1061 AspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGluGlyLeuPro 1080  
DB 3214 GATGAGATCAAGAAGCGCAATGTCAGCAAAATATCCATTTCTTTATGAGGTCTCCCT 3273  
QY 1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 1100

DB 3274 GAGAAATACAAACACCAAGTTGGAGCTGAAAGGAGGACACGCTTTCTGGCGCCAGAAACAA 3333  
QY 1101 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120  
DB 3334 AGACTAGCTATTGCAAGGCTCTTCTCCAAAACCCAAAATTTTATTTGTTGGATGAGGCC 3393  
QY 1121 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140  
DB 3394 ACTTCAGCCCTCGATAATGACAGTGAGAAAGTGGTTTCAGCATGCCCTTGATAAAGCCAGG 3453  
QY 1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160  
DB 3454 ACGGGGAAGACATGCTAGTGTCTCAAGGCTCTCTGCAATTTCAGAACCGCAGATTG 3513  
QY 1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180  
DB 3514 ATATGGTCTTCGCACANATGGAAGATTAAGGAACAAGGAACCTCATCAAGAGCTCTTGAGA 3573  
QY 1181 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195  
DB 3574 AATCGAGACATATATTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3618

RESULT 2

US-09-873-409-13  
; Sequence 13, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 13  
; LENGTH: 3702  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: NO. US20020037522A1e  
; LOCATION: (723)..(723)  
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)  
US-09-873-409-13

Alignment Scores:  
Pred. No.: 0 Length: 3702  
Score: 5988.50 Matches: 1195  
Percent Similarity: 97.79% Conservative: 0  
Best Local Similarity: 97.79% Mismatches: 0  
Query Match: 99.61% Indels: 27  
DB: 10 Gaps: 1

US-09-873-409-6 (1-1195) x US-09-873-409-13 (1-3702)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 34 ATGATCTCGGTATATCTGGCATCACTGGTCAATGGAGCCCTTCCTTTAATGCCACTG 93  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 94 GTTTTAGGAGAAATGATGATAACCTTATTAGTGGATGTCTAGTCCAAACCAACATAC 153  
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
DB 154 TCTTCTTCAGGTGACCCCTGATTATGTTGGAATAGGTGTTGCTGCTTGAATTTTGGT 213  
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
DB 214 TACATACAGATTTTCTTGTGGATTATTAACGTGCAGCAGACAGAGGATTCGAAAA 273

QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
DB 274 CAGTTTTTTCATTTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333  
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 334 GAACCTTAACACTCGCATGCAGACATTCACAAATCAGTGTATGAGATAAGATT 393  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys 140  
DB 394 GCCTGTGTGTTTCAAAACATGCTACTTTTTTCGATTCGGCTGGCAGTTGTTGGTGAAG 453  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 454 GGCTGGAACTCACCTAGTGACTCTATCCAGCTCTCTCTTATATGGCTTCAGCGGCA 513  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaThrSerLysAla 180  
DB 514 GCATGTGTTCTAGGATGGTCACTCATTCAGCAGTAAGGAATTAAGTGCTTATTCCAAAGCT 573  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCCTTTAGGGCCAG 633  
QY 201 GluLysGluLeuGln----- 205  
DB 634 GAGAAAGAACTTCAAAGGCTCTTCCCTTTTAAATATAACAAGATATGCTTGGTTTATTTT 693  
QY 206 -----ArgTyrThrGlnAsnLeuLysAsp 213  
DB 694 CCCAGTGGCTACTAAGTGTGTTCTGTTTGTAAAGGTATACACAGAATCTCAAAGAT 753  
QY 214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233  
DB 754 GCAAGGATTTTGGCAATAAAGAGACTATAGCTTCAAAAGTGTCTCTGTGTGTGTGTAC 813  
QY 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253  
DB 814 TTCTTTATGATGGAACCTATAGACTTGTCTTTTGGTATGGAACCTCTTGTATCTTAAT 873  
QY 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273  
DB 874 GGAAACCTCGATATACATCGGAGCTGTTCTGCTGTTTCTTTAGTGTAAATCCATAGC 933  
QY 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293  
DB 934 AGTTATTTGCATTTGGACGAGCAGTCCCTCCTTGAACCTTCGCAATAGCCCGAGGAGCT 993  
QY 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313  
DB 994 GCCTTTTCATATTTCCAGGTATTTGATAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053  
QY 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333  
DB 1054 GGAATATAACCTGAATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTCAATAT 1113  
QY 334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353  
DB 1114 CCATCAAGACCATCTATCAAGATTTCTGAAGGCTCGAATCTCAAGATTAAGTCTGGAGAG 1173  
QY 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373  
DB 1174 ACATCGCCTTGGTTCGCTCTCAATGGCAGTGGGAAGTAGTACGGTAGTTCAGCTTTCGAG 1233  
QY 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393  
DB 1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293  
QY 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413  
DB 1294 AATGTGGGCAATATTCAGACCAATTTTGGAGTGGTTAGTCAAGAGCCCTGTTTGTTCGGG 1353  
QY 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 433

DB 1354 ACCACCATCAGTAACAATATCAAGTATGACGAGATGATGTGCTGATGAAGAGATGGAG 1413  
QY 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453  
DB 1414 AGAGCAGCAAGGGAAGCAAAATCGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473  
QY 454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473  
DB 1474 ACATTTGGTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGCGCAAAACAGAGGATCGCAAT 1533  
QY 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493  
DB 1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCCTGATTTTAGATCAGGCTACGCTGCGCTG 1593  
QY 494 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThr 513  
DB 1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACCTGGAGAAAGCGCAAGGTCGGACT 1653  
QY 514 ThrIleValAlaAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 533  
DB 1654 ACNATCGTGTAGCACACCGACTTCTACTATTCGAAAGTGCAGATTTGATTTGACCCCTA 1713  
QY 534 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 553  
DB 1714 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACCTAATGGCAAAACGAGGTCTA 1773  
QY 554 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 573  
DB 1774 TATATTTCATCTGTGATGTGCACAGGATATTTAAAGCTGTGATGAACAGATGGAGTCAATG 1833  
QY 574 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 593  
DB 1834 ACNATTTCTACTGAAAGAAAGACCAACTCACTCTCTGCACCTCTGTGAAGCATCAAG 1893  
QY 594 SerAspPheIleAspLysAlaGluSerThrGlnSerLysGluLysSerLeuProGlu 613  
DB 1894 TCAGACTTCATTGACAAGCTGAGGAATCCACCAATCTAAAGAGATAAGCTTCCTGAA 1953  
QY 614 ValSerLeuLeuLysIleLysLeuAsnLysProGluTrpProPheValValLeuGly 633  
DB 1954 GTCTCTCTATTAATAAATTTTAAAGTTAAACCAAGCTGAAATGGCCTTTTGTGGTCTGGGG 2013  
QY 634 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 653  
DB 2014 ACATTTGGCTTCTGTTCTTAATGGAACCTGTTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073  
QY 654 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 673  
DB 2074 ATTATTAACCATGTTTGGAAAAATATGATAAAACCAATTAAGCATGATGCAGAAATTTAT 2133  
QY 674 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 693  
DB 2134 TCCATGATATTCGTCAATTTGGGTGTATTTGCTTGTGCTAGTTATTTTCATGCGAGGATTA 2193  
QY 694 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 713  
DB 2194 TTTTACGGCAGACAGCGGGAATTTTAACATGAGATTAAGACACTTGGCCTTCAAGGCC 2253  
QY 714 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 733  
DB 2254 ATGTTATATCAGGATATTCCTGCTGTTTGTATGAAAGGAAAAACAGCACAGGAGCTTGACA 2313  
QY 734 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 753  
DB 2314 ACAATATTAGCCCATAGATATAGCACAAATTCAGAGGAGCAACAGGTTCAGGATTTGGCGTC 2373  
QY 754 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 773  
DB 2374 TTAACCAAAATCCAACTAACATCGGACTTTCAGTTATCATTTCTTTATATATGATGG 2433  
QY 774 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 793

```
Db 2434 GAGATGACATTCCTCGATTCTGAGTATTGCTCCAGTACTTCCCGTGACAGCAATGATTGAA 2493
Qy 794 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 813
Db 2494 ACCGACGCAATGACTGGATTTCGCAACAAGATAAGCAAGAACTTAAGCATGCTGGAAG 2553
Qy 814 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 833
Db 2554 ATAGCAACTGAAAGCTTTGGAGAAATATACGTACTATATAGTGTCTATTAAACAAGGGGAAAAGCC 2613
Qy 834 PheGluGlnMetThrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 853
Db 2614 TTCGACCAATGATGAGAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAGCA 2673
Qy 854 GlnIleGlySerCysTyrrAlaPheSerHisAlaPheIleTyrrPheAlaTyrrAlaAla 873
Db 2674 CAGATTATTGGAAGCTGTTATGCAATTCAGCCATGCGCTTTATATATATTTCCTATGCGCA 2733
Qy 874 GlyPheArgPheGlyAlaTyrrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 893
Db 2734 GGGTTTCGATTTCGAGCCTATTTAATTCAGCTGGACGAATGACCCACAGAGGCATGTTTC 2793
Qy 894 IleValPheThrAlaIleAlaTyrrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
Db 2794 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2853
Qy 914 ProGluTyrrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLys 933
Db 2854 CCTGATATTCCAAAGCCAAATCGGGGCTGCGCATCTGTTGCTTGTGGAAAAGAAA 2913
Qy 934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953
Db 2914 CCAATATATAGACGCGCAGTCAAGAAGGAGAAAGCCAGACACATGTGAAGGGAATTTA 2973
Qy 954 GluPheArgGluValSerPhePheTyrrProCysArgProAspValPheIleLeuArgGly 973
Db 2974 GAGTTTCGAGAAGTCTCTTTCTTCATCCATGTCGCCAGATGTTTTCATCTCCCGTGGC 3033
Qy 974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
Db 3034 TTATCCTCTAGTATTGAGCGAGGAGAAAGACAGTAGCATTTGTGGGGAGACGCGCTGGG 3093
Qy 994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrrAspProValGlnGlyGlnValLeu 1013
Db 3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTCGAAGGACAAAGTGCTG 3153
Qy 1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033
Db 3154 TTTGATGGTGTGGATGCAAAAAGAAATTGAAATGTATACAGTGGCTCCGTTCCCAAAATAGCAATC 3213
Qy 1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrrGlyAsp 1053
Db 3214 GTTCTCTAAGAGCCTGTGCTCTTCAACTGCAGCATTTGCTGAGAACATCGCCTATGTTGAC 3273
Qy 1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1073
Db 3274 AACAGCGGTGTGTGCAATTAGATGAGTCAAGAAGCGCAAAATGCAGCAAAATATCCAT 3333
Qy 1074 SerPheIleGluGlyLeuProGluLysTyrrAsnThrGlnValGlyLysGlyAlaGln 1093
Db 3334 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGGAAGAGACACAG 3393
Qy 1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLys 1113
Db 3394 CTTTCTGGCGGCCAGAAAACAAGACTAGTATTGCAAGGGCTCTTCTCCAAAACCCAAA 3453
Qy 1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
Db 3454 ATTTTATTGTTGATGAGGCCACTTCAGCCCTCGATPATGACAGTGAAGGTTGTTTCAG 3513
Qy 1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
Db 3514 CATGCCCTTGATAAAGCCAGGACGGGAAGACATGCTAGTGGTCACTCACAGGCTCTCT 3573
```

```
Qy 1154 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1173
Db 3574 GCAATTTCAGAACGCAGATTGTAGTGGTTCTGCACAATGGAAGATAAAGGAACAAGGA 3633
Qy 1174 ThrHisGlnGlnLeuLeuArgAsnArgaspIleTyrrPheLysLeuValAsnAlaGlnSer 1193
Db 3634 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTTAGTGAATGCACAGTCA 3693
Qy 1194 ValGln 1195
Db 3694 GTGCAG 3699
RESULT 3
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12
Alignment Scores:
Pred. No.: 0 Length: 3177
Score: 5159.50 Matches: 1031
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 0
Query Match: 85.82% Indels: 27
DB: 10 Gaps: 1
US-09-873-409-6 (1-1195) x US-09-873-409-12 (1-3177)
Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCATCTCATTTGACCAGTAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGCTGTGGCA 60
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGTCTTGTCATCAATCCGAACGATAGCTTTAGGGCCCCAGGAGAAAGAACTT 120
Qy 205 Gln----- 205
Db 121 CAAAGTCTTCTTTAAATATAACAAGATATGCTTGGTTTATTTTCCCCAGTGGCTA 180
Qy 206 -----ArgTyrrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 181 CTAAGTTGTGTTCTGTTNTTTTGAAGGTATACAGAAATCTCAAGATGCAAAAGGATTTT 240
Qy 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrrPhePheMetAsn 237
Db 241 GGCATAAAAGGACTATAGCTCAAAAAGTGTCTCTGGTGTGTGTACTTCTTTTATGAAT 300
Qy 238 GlyThrTyrrGlyLeuAlaPheTrpTyrrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
Db 301 GGAACCTATGGACTTGTCTTTTGGTATGGAACCTCCTCGATCTTAATGGAGAACCTGGA 360
Qy 258 TyrrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrrCysIle 277
```

361	Db	TATACCAATCGGAGACTGTTCTTGCTGTTTCTTTAGTGTAAATCCATAGCAGGTATTGCATT	420
278	Qy	GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle	297
421	Db	GGAGCAGCAGTCCCTCACTTTGAACCTTTCCGAANTAGCCCGAGGAGCTGCCTTTTCATATT	480
298	Qy	PheGlnValIleAspIlyAspProSerIleAspAsnPheSerThrAlaGlyTyrLysPro	317
481	Db	TTCCAGGTATTATGATAAGAAACCCAGTATAGATAAATCTTTCCACAGCTGGATATAAACCT	540
318	Qy	GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro	337
541	Db	GAATTCCTATAGAGAACTGTGGAAATTTAAAAATGTTCTTTTCAATATTCATCAAGACCA	600
338	Qy	SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu	357
601	Db	TCATATCAAGATTCTGAAGGTCCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCTTG	660
358	Qy	ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp	377
661	Db	GTCCGCTCTCAATGCCAGTGGGAAGAGTAGCGTAGTCCAGCTTCTGCACAGGTTATATGAT	720
378	Qy	ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis	397
721	Db	CCCGATGATCGCTTATCATGTGTGGATGAGAATGCATCAGAGCTTTAAATGTGCGGCAT	780
398	Qy	TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer	417
781	Db	TATCGACACCATATTTGGAGTGGTTAGTCAAGAGCCCTGTTTTGTTCCGGACCCCATCAGT	840
418	Qy	AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg	437
841	Db	AACAATATCAAGTATGGACGAGATGATGTGCTGATGAAGAGATGGAGAGCAGCAAGG	900
438	Qy	GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly	457
901	Db	GAAGCAAAATCGTATGATTTTATCATGCGAGTTTTCCTAATAAATTTAATACATTCGTAGGG	960
458	Qy	GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAatqAlaLeu	477
961	Db	GAATAAGGAGCTCAATATGATGTGAGGGCGAAGAACAGAGATCGCATTTGCTGTGCGCTTA	1020
478	Qy	ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer	497
1021	Db	GTTCGAACCCCAAGATTCTGATTTTATGATGAGGCTACGTCTGCCTCGATTACAGAAAGC	1080
498	Qy	LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal	517
1081	Db	AAGTCAGCTGTTCAAAGTGCACCTGGAGAAGGCGCAGCAAGAGTCGAGCTACCAATCGTGGTA	1140
518	Qy	AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet	537
1141	Db	GCACACCGACTTCTACTATTCGAAGTCGAGATTTGATTGTGACCTTAAGAGATGGAAATG	1200
538	Qy	LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu	557
1201	Db	CTGCGCGAGAAAGGACACATGCTGAACCTAATGGCAAAACGAGGTCTATATATTCACTT	1260
558	Qy	ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr	577
1261	Db	GTGATGTCACAGGATATTAATAAAGCTGATGAACAGATGGAGTCAATGCACATATTCCTACT	1320
578	Qy	GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle	597
1321	Db	GAAGAAGAAGCAACCTCACCTTCCTCGCACTCTGTGGAAGAGCATCAAGTCAGACTTCATT	1380
598	Qy	AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu	617
1381	Db	GACAAGGCTCAGGAATCCACCCCAATCTAAGAGATAAGTCTTCTCGAAGTCTCTCTATTATA	1440
618	Qy	LysIleLeuLysLeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSer	637
1441	Db	AAAAATTTTAAAGTTAAACAGACCTGAATGGCCCTTTTGTGTTCTCTGGGACATCGGCTCT	1500

Qy	638	ValleuAenGlyThrVallHisProValPheSerIleIlePheAlaLysIleIleThrMet	657
Db	1501	GTCTCAAATGGAACTGTTTCATCCAGTATTTTCCATCATCTTTTGGAAAAATTTATAACCATG	1560
Qy	658	PheGlyAenAenAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe	677
Db	1551	TTTGGAAATTAATGATTAACCAACATTTAAAGCATGATGAGAAAATTTATTCATGATATTC	1620
Qy	678	ValleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg	697
Db	1621	GTCAATTTGGGTGTATTTGCTTTGTGCATTAATTTCAATGAGGAGTATTTTACGGCAGA	1680
Qy	698	AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln	717
Db	1691	CGAGGGAAATTTTAAAGATGAGATTAAGACACTTGGCCCTTCAAGAGCCATGTTATATCAG	1740
Qy	718	AspIleAlaTrpPheAspGluLysGluAenSerThrGlyGlyLeuThrThrIleLeuAla	737
Db	1741	GATATTGCCTGGTTTGTATGAAAGGAAAAACACACAGAGAGGCTTGACAACAATATTAGCC	1800
Qy	738	IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAen	757
Db	1801	ATAGATATAGCACAAATTTCAAGGAGCAACAGAGTTCAGAGATTGGCGCTTTAAACCAAAAT	1860
Qy	758	AlaThrAenMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe	777
Db	1861	GCAACTAACATGGACCTTCAGTTATCAATTCCTTTATATATGATGGAGAGATCATTC	1920
Qy	778	LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet	797
Db	1921	CTGATTCTGAGTATTGCTCCAGTACTTGCCTGAGCAGAGAAATGATTGAAACCCGACAAATG	1980
Qy	798	ThrGlyPheAlaAenLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu	817
Db	1981	ACTGGATTGCCAACAAAGATAGCAAGAACTTTAAGCATGCTTGGAAAGATAGCAACTGAA	2040
Qy	818	AlaLeuGluAenIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet	837
Db	2041	GCCTTTGGAGAAATACGTACTATAGTGTCAATTAACAAGGAAAAAGCCCTTCGAGCAAAATG	2100
Qy	838	TyrGluGluMetLeuGlnThrGlnHisArgAenThrSerLysLysAlaGlnIleIleGly	857
Db	2101	TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA	2160
Qy	858	SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe	877
Db	2161	AGCTGTTATGCATTACGCCATGCTTTATATATTTTGCTATGACAGAGGGTTTCGATTT	2220
Qy	878	GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr	897
Db	2221	GGAGCCTATTAAATTCAAAGCTGACGAAATGACCCACAGAGGGCATCTTCATAGTTTTTACT	2280
Qy	898	AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer	917
Db	2281	GCAATTGCATGAGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCGTGAATATTTC	2340
Qy	918	LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp	937
Db	2341	AAAGCCAAATCGGGGCTGGCATCTGTTTGGCTTCTTGGAAAAACCAAAATATAGAC	2400
Qy	938	SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAenLeuGluPheArgGlu	957
Db	2401	AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAAATTTAGAGTTTCGAAA	2460
Qy	958	ValSerPheThrTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer	977
Db	2461	GTCTCTCTTCTATATCATGTGCCAGATGTTTTCATCTCTGGTGGCTTATTCCTCAAT	2520
Qy	978	IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer	997
Db	2521	ATTGAGCGAGAAAGACAGTAGCATTTGTGGGAGCAGCGGCTGTGGGAAAGACACTTCT	2580



Qy	998	ValGlnLeuLeuGlnArgLeuTyfAspProValGlnGlnGlyClnValLeuPheAspGlyVal	1017
Db	2581	GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAGAGGACAAGTGTGTTGATGCTGTG	2640
Qy	1018	AspAlaLysGluLeuLeuValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu	1037
Db	2641	GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCTCAATAGCAATCGTTCCTCAAGAG	2700
Qy	1038	ProValLeuPheAsnCyssSerIleAlaGluAsnIleAlaTyfGlyAspAsnSerArgVal	1057
Db	2701	CCTGTGCTCTTCAACTGCAGCAATTCGTGAGAACATCGCCTATGTCACCAACAGCGGTG	2760
Qy	1058	ValProLeuAspGluIleLysGluAlaAlaAsnAlaAsnIleHisSerPheIleGlu	1077
Db	2761	GTGCCATTAGATGAGATCAAGAAGCGCAAAATGCAGCAAAATATCCATTCTTTTATTGAA	2820
Qy	1078	GlyLeuProGluLysTyfRanThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly	1097
Db	2821	GGTCTCCCTGAGAAATACAAACACAAAGTTGGACTGAAAGGACACAGCTTTCTGCGGC	2880
Qy	1098	GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu	1117
Db	2881	CAGAAACAAGACTAGCTATTGCAAGGCTCTTCTCCAAAACCCCAAAATTTTATTGTTG	2940
Qy	1118	AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp	1137
Db	2941	GATGAGGCCACTTCAGCCCTCGATATGACAGTGAGAAAGGTGTTTCAGCATGCCCTTGAT	3000
Qy	1138	LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn	1157
Db	3001	AAAGCCAGGACGGGAGGACATCCCTAGTGGTCACACAGGCTCTCTGCAATTCAGAAC	3060
Qy	1158	AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGlu	1177
Db	3061	GCAGATTGTAGTAGTGGTCTGCAATGGAAGTAAAGGAAACAAGAACTCATCAAGAG	3120
Qy	1178	LeuLeuArgAsnArgAspIleTyfPheLysLeuValAsnAlaGlnSerValGln	1195
Db	3121	CTCCTGAGAAATCGACATATATTTTAAAGTTAGTGAATGCACAGTCAGTGCAG	3174
RESULT 4			
US-09-873-409-10			
; Sequence 10, Application US/09873409			
; Patent No. US20020037522A1			
; GENERAL INFORMATION:			
; APPLICANT: Frank, Markus			
; APPLICANT: Sayegh, Mohamed			
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein			
; FILE REFERENCE: 81994/268611			
; CURRENT APPLICATION NUMBER: US/09/873,409			
; CURRENT FILING DATE: 2001-06-05			
; NUMBER OF SEQ ID NOS: 19			
; SOFTWARE: Patentin version 3.0			
; SEQ ID NO 10			
; LENGTH: 2856			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-873-409-10			
Alignment Scores:			
Pred. No.:	0	Length:	2856
Score:	4561.50	Matches:	912
Percent Similarity:	97.96%	Conservative:	0
Best Local Similarity:	97.96%	Mismatches:	1
Query Match:	75.87%	Indels:	18
DB:	10	Gaps:	1
US-09-873-409-6 (1-1195) x US-09-873-409-10 (1-2856)			
Qy	266	ValPhePheSerValIleHisSerSerTyfCysIleGlyAlaAlaValProHis-PheGI	285
Db	114	GTTTCTTTAGTGTAAATCCATAGCAGTTATTGCAATTGAGCAGCAGTCCCTCATTTATT--	171

Qy	285	uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPr	305
Db	172	-----GATAAGAAACC	182
Qy	305	oSerIleAspAenPheSerThrAlaGlyTyfLysProGluSerIleGluGlyThrValGI	325
Db	183	CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGGA	242
Qy	325	uPheLysAenValSerPheAsnTyfProSerArgProSerIleLysIleLeuLysGlyLe	345
Db	243	ATTTAAATATGTTCTTTCATATTCATCAAGACCATCTATCAGATTCTGAAAGTCT	302
Qy	345	uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy	365
Db	303	GAATCTCAGAAATAAGTCTGGAGAGACAGTCGCTTGGTCGGTCTCAATGGCAGTCGGA	362
Qy	365	sSerThrValValGlnLeuGlnArgLeuTyfRaspProAspAspGlyPheIleMetVa	385
Db	363	GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT	422
Qy	385	lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyfArgAspHisIleGlyValVa	405
Db	423	GGATGAGATGACATCAGAGCTTTAAATGTCGGCATTTATCGAGACCATATTGGAGTGT	482
Qy	405	lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyfGlyArgAs	425
Db	483	TAGTCAAGAGCCTGTTTGTTCGGACCAACCATCAGTAACAATATCAAGTATGGACGAGA	542
Qy	425	phasValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyfRaspPheI	445
Db	543	TGATGTGACTGATGAAGAGATCGAGAGACGACGCAAGGGAAGCAAAATGCGTATGATTTAT	602
Qy	445	eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGI	465
Db	603	CATGGAGTTTCTTAATAATTTAATACATTTGTAGGGGAAAAGAGGCTCAATATGAGTGG	662
Qy	465	YGIglnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuI	485
Db	663	AGGCGAGAAACAGAGGATCGCAATTCCTGCTTCCTTAGTTTCAAAACCCCAAGATTCTGAT	722
Qy	485	eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaIle	505
Db	723	TTTAGATGAGGCTACGTCCTGCTCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAGACTC	782
Qy	505	uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr	525
Db	783	GGAGAGGCGAGCAAGGTCGGACTACATTCGTTGTTAGCACACCCGACTTTCTACTATTTC	842
Qy	525	gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl	545
Db	843	AAAGTCAGATTTGATTTGACCCCTAAAGGATGGAATGCTGGCGGAGAAAGGACACATGC	902
Qy	545	aGluLeuMetAlaLysArgGlyLeuTyfRsrLeuValMetSerGlnAspIleLysLy	565
Db	903	TGAACCTAATGGCAAAACGAGGCTATATTTATTTACCTTGTGATGTCACAGGATATTA	962
Qy	565	sAlaAspGluGlnMetGluSerMetThrTyfSerThrGluArgLysThrAsnSerLeuPr	585
Db	963	AGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAGAAAGAACCCACTCATTC	1022
Qy	585	oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGI	605
Db	1023	TCTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCAATGACAGGCTGAGGAATCCACCCA	1082
Qy	605	nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr	625
Db	1083	ATCTAAAGAGATAAGTCTTCTCCTGAAAGTCTCTATTTAAATAATTTTAAAGTTAAACAG	1142
Qy	625	oGluTrpProPheValIleGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr	645
Db	1143	TGAATGCCTTTTGTGGTCTCTGGGACATTTGGCTTCTTTCTTAAATGGAACTGTTTCATCC	1202

645 oValPheSerIleIlePheAlaIysIleIleThrMetPheGlyAsnAsnAspIysThrTh 665  
1203 AGTATTTTCCATCATCTTTGCAAAATTAATACCATGTTTGGAAATAATGATAAAACCAC 1262  
665 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 695  
1263 ATTAAGCATGATGCAGAAATTTATTCATCATGATATTCGTATTTGGGTGTTATTTGCTT 1322  
685 eValSerTyrPheMetGlnGlyLeuPheTyrGlyAraAlaGlyGluIleLeuThrMetAr 705  
1323 TGTGAGTTATTTTCATGAGGATTTATTTTACGGCAGACGAGGGAATTTTAAACGATGAG 1382  
705 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLy 725  
1383 ATTAAGACACTTGGCTTCAAAGCATGTTATATATCAGGATATTGCCTGGTTTGTATGAAA 1442  
725 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnI 745  
1443 GGAACACAGCACAGGAGCTTGACAACAATATTAGCCATAGATATAGCAACAATTTCAAGG 1502  
745 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 765  
1503 AGCAACAGGTTCCAGGATTTGGCTTTAAACACAAATGCAAACTAACATGGGACTTTTCAGT 1562  
765 lIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVa 785  
1563 TATCATTTCTTATATATGATGGAGATGACATTCCTGATTCGATTTGCTCCAGT 1622  
785 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 805  
1623 ACTTGGCTGACAGCAATGATTAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682  
805 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrI 825  
1683 GCAAGAACTTAAGCATCTGGAAGATAGCAACTGGAAGCTTTGGAGAAATATACGTACTAT 1742  
825 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 845  
1743 AGTGCTATTAAACAGGAGAAAGCTTCGAGCAATGATTAAGAGATGCTTCAGACTCA 1802  
845 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 865  
1803 ACACAGAAATACCTCGAAGAAAGCACAGATATTATGGAAGCTGTTATGCAATTCAGCCATGC 1862  
865 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleLeuGlnAlaGl 885  
1863 CTTTATATATTTTCCCTATGCAGCAGGTTTCGATTTTGGAGCTTATTTAATTTCAAGCTGG 1922  
885 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 905  
1923 ACGAATGACCCAGAGGCGATGTTTCATAGTTTTCATGCAATTCATATGGAGCTATGGC 1982  
905 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 925  
1983 CATCGAAAAACGCTCGTTTTGGCTCCTCGAATATTCAAAGCCAAATCGGGGGCTGGCA 2042  
925 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 945  
2043 TCTGTTTGGCTTGTGGAAAAAGAAACCAAAATATAGACGCGCAGTCAAGAAAGGAAAAA 2102  
945 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 965  
2103 GCCAGACACATGTGAGGGAATTTAGAGTTTCGAGAGTCTCTTCTTCTATCCATGCTCG 2162  
965 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 985  
2163 CCCAGATGTTTTCATCTCCGCTGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222  
985 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTy 1005  
2223 ATTTGTGGGAGACGCGCTGTGGAAAAAGCACTTCTGTTCACACTTCTGCAGAGACTTTA 2282  
1005 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1025

2283 TGACCCCGTCGAAGGACAGTCTGTTGATGCTGTGGATGCAAAAGAAATTGAATGTACA 2342  
1025 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1045  
2343 GTGGCTCCGTTTCCAAAATAGCAATCGTTCTCAAGAGCCTGTCTTCACTGACGAT 2402  
1045 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1065  
2403 TGCTGAGACATCGCTATGTTGTCACACAGCCGTTGGTGGCCATTAGATGAGATCAAAGA 2462  
1065 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1085  
2463 AGCCGCAATGACGCAAAATATCCATTTTATGAAGTCTCTCCCTGAGAAAATACAAACAC 2522  
1085 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnAtcLeuAlaIleAl 1105  
2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGCCAGAAACAAAGACTAGCTATTGC 2582  
1105 aArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAs 1125  
2583 AAGGCTCTTCTCCAAAACCCCAAAATTTATTTGTTGATGAGGCCACTTCAGCCCTCGA 2642  
1125 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1145  
2643 TAATGACAGTGAGAGGTGGTTTCAGCATGCCCTTGATAAGCCAGGACGGAAGGACATG 2702  
1145 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1165  
2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTGAGAACGAGACTTTGATAGTGTCTGCA 2762  
1165 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuAsnArgAspIleTy 1185  
2763 CAATGGAAGATTAAGAAACAAAGAACTCATCAAGAGCTCTCTGAGAAATCGAGACATATA 2822  
1185 rPheLysLeuValAsnAlaGlnSerValGln 1195  
2823 TTTTAAGTTAGTGAATGCACAGTCAGTCAGCAG 2853

## RESULT 5

US-09-866-866A-1  
; Sequence 1, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 3860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-866-866A-1

Alignment Scores:  
Pred. No.: 7,94e-311 Length: 3860  
Score: 3554.00 Matches: 682  
Percent Similarity: 75.63% Conservative: 246  
Best Local Similarity: 55.58% Mismatches: 263  
Query Match: 19.12% Indels: 36  
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-866-866A-1 (1-3860)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu	20
Db	151	ATGTGGTGGCAACTTTGGCTGCCATCATCCATCGGGCTGGACCTTCCTCTCATGATCGTG	210
Qy	21	ValLeuGlyGlumetSerAsp-----AenLeuIleSer	31
Db	211	GTGTTGGAGAAATGACAGATACTTTGCMAATGAGGAATTTAGAAGATCTCATGTCA	270
Qy	32	GlyCysLeuValGlnThr-----AenThyrSerPhePhe-----	43
Db	271	AACATCACTAATAGAAGTGATCAATGATACAGGGTTCCTCATGAATCTGGAGGAAGAC	330
Qy	44	-----ArgLeuThrLeuThyrTyrrValGlyIleGlyValAlaLaLeuIlePheGlyTyr	61
Db	331	ATGACCAGATATCGCTATTATTACAGTGGNAITGGTGTGGGTGCTGGTTGCTGCTTAC	390
Qy	62	IleGlnIleSerLeuThrpIleIleThrAlaAlaArgGlnThrLysArgIleAtrGLvsGln	81
Db	391	ATTCAAGTTTTCTTTGGTGCTCGCAGCTGGAGACAATAACAAAAATTAGAAAACAG	450
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101
Db	451	TTTTTTCATGCTAATCGCACAGGAGATAGCTGGTTGATGTGCACGATGTTGGGGAG	510
Qy	102	LeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	511	CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGGTATTGGTGACAAAATT	570
Qy	121	AlaLeuLeuPheGlnAenMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	571	GGAAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAGT	630
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	631	GGTTGGAAGCTAACCCCTTGATGATTTTGGCCATCAGTCCTGTTCTTGGACTGCTCAGCTGT	690
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	691	GTCTGGGCAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGCTATGCAAAAAGCT	750
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	751	GGACAGTAGCTGAGAGGCTCTGGCAGCAATTAGNACTGTGATGTCATTTGGAGGCAA	810
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAenLeuLysAspAlaLysAspPheGlyIleLys	220
Db	811	AAGAAAGAACTTGAAAGGTACAAACAAAAATTTAGAAGAAAGCTTAAAGAAATTTGGGATAAG	870
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	871	AAAGCTATTACAGCCAATATTCTATAGGTGCTGCTTCTCGCTGATCTATGCAATCTTAT	930
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAenGlyGluProGlyTyrThrIle	260
Db	931	GCTCTGGCCTCTCGTATGGACACACCTTGGTCTCTCAGGGGAA-----TATTCTATT	984
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	985	GGACAAGTACTCACTGTATTCTTTTCTGTATTAAATGGAGGAGAGCTATTAGTAATCTTCAAGATA	1044
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	1045	TCTCCAAGCATTTGAAGCAITTTGCAAAATGCAAGGAGAGAGCTATTAGAAATCTTCAAGATA	1104
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1105	ATTGATAAAGCCAACTATTGACAGCTATTGCAAGAGTGGGCACAAACCAAGATAAATT	1164
Qy	321	GluGlyThrValGluPheLysAenValSerPheAsnTyrProSerArgProSerIleLys	340
Db	1165	AAGGGAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAAGTTAAG	1224

Qy	341	IleLeuLysGlyLeuAsnLeuArgLleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1225	ATCTTTGAAGGCCCTGAACCTGAAGTGGCAGAGTGGCGACGGTGCCTCGTTGGAAAC	1284
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyAspProAspAsp	380
Db	1285	AGTGGCTGTGGGAAGACACACAGTCCAGCTGTATGCAGAGGCTCTATGACCCACAGAG	1344
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp	400
Db	1345	GGGATGGTCAGTGTGTGATGCACAGGATATTAGGACCATAAATGTAAGGTTCCTACCGGAA	1404
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1405	ATCATTTGGTGTGGTACGAGGAACCTGTATTGTTTGGCCACACAGATGATGAACAACAT	1464
Qy	421	LysTyArgLysArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1465	CGCTATGGCCGTGAATAATGTCCATGTGATGAGATTGAGAAAGCTGTCCAGGAGCCCAAT	1524
Qy	441	AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly	460
Db	1525	GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCTCGTTGGAGAGAGGG	1584
Qy	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1585	GCCCAAGTTGAGTGGTGGCCAGAACGACAGAGATCGCCATTGTCACGTCCTCGTTGCGAAC	1644
Qy	481	ProLysIleIleuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1645	CCCAAGATCTCTCTGGGATGAGGCCAGCTCAGCCTTGGACACAGAACGACGACGTG	1704
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1705	GTTCAAGTGGCTCTGGATAAGCCAGAAAGGTCGACACCATTTGTGATAGCTCATCGT	1764
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1765	TTGTCTACAGTTCGTAATGCTGACATCGCTATCGCTGGTTTCGATGATGGAGTCAATTGGAG	1824
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSer	560
Db	1825	AAAGAAATCATGATGAACCTCATGAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG	1884
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	1885	CACACAGCAGCAAAATGAAGTTGAATTAGAAAGATGCAGCTGATGAATCCAAAGTGAAT	1944
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	1945	GATGCCITTTGGAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT	2004
Qy	590	---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys	607
Db	2005	CGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAGAGAGCTCTG	2066
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp	627
Db	2065	GATGAAGATATACCTCCAGTTCCTTTTGGAGGATTATGAAGCTAAATTAATTAACGATGG	2124
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2125	CCCTATTTTTGTGGTGTATTTTGTGCCATTATAATGGAGGCCCTGCAACACGACATTT	2184
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2185	GCAATPAATATTTTCAAAGATATATAGGGGTTTTTACAAAGAAATTTGATGATCCTCGAAACAAA	2244
Qy	667	LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2245	CGACAGAAATGATGATGTTTTCATATTGTTTCTAGCCCTTGGAAATTAATTTCTTTTAT	2304
Qy	687	SerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyIleLeuLeuThrMetArgLeu	706

```
Db 2305 ACATTTTCTCTCAAGGTTTCACATTTGGCAAGCTGGAGAGATCTCTCAACCAAGCGGCTC 2364
      :::::::::::::::::::::
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaTrpPheAspGluLysGlu 726
      :::::::::::::::::::::
Db 2365 CGATACATGGTTTCCGATCCATGCTCAGACAGAGATGGATGGTTGGTTGATGACCCCTAAA 2424
      :::::::::::::::::::::
Qy 727 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
      :::::::::::::::::::::
Db 2425 AACACCACTGGAGCATTGACTACAGGCTGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484
      :::::::::::::::::::::
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766
      :::::::::::::::::::::
Db 2485 ATAGGTTCCAGGCTGTGTAATTTACCCAGATATATAGCAAAATCTTGGCAGCAATAAAT 2544
      :::::::::::::::::::::
Qy 767 IleSerPheIleTyGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
      :::::::::::::::::::::
Db 2545 ATATCCTTTCATCTATGTTGGCAACTAACACTGTACTTACCAATTTGACCCATCAT 2604
      :::::::::::::::::::::
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaLeuLysAspLysGln 806
      :::::::::::::::::::::
Db 2605 GCATACACAGAGTGTGTGAATGAAATGTTGCTGCGACACGACTGGAAGATAGAAA 2664
      :::::::::::::::::::::
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
      :::::::::::::::::::::
Db 2665 GAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAAACCGTTGTT 2724
      :::::::::::::::::::::
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeuGlnThrGlnHis 846
      :::::::::::::::::::::
Db 2725 TCTTTGACTCAGGAGCAGAGTGTGAACATATGTATGCTCAGAGTTTTCAGGTTACCAATAC 2784
      :::::::::::::::::::::
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyRAlaPheSerHisAlaPhe 866
      :::::::::::::::::::::
Db 2785 AGAACTCTTTGAGAAAGCACACATCTTTGGAATTACATTTTCCTTCAACCCAGGCAATG 2844
      :::::::::::::::::::::
Qy 867 IleTyRPhaIatyrAlaAlaGlyPheArgPheGlyAlatyrLeuIleGlnAlaGlyArg 886
      :::::::::::::::::::::
Db 2845 ATGTATTTTCTCTATGCTGGATGTTTCGGGTTTGAGCCCTACTTGGTGCCACATAAACTC 2904
      :::::::::::::::::::::
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyRAlaMetAlaIle 906
      :::::::::::::::::::::
Db 2905 ATGAGCTTTGAGGATGCTCTGTAGTATTTTTCAGCTGTTGCTTGGTGCATGGCCGCTG 2964
      :::::::::::::::::::::
Qy 907 GlyLysThrLeuValLeuAlaProGluTyRSerLysAlaLysSerGlyAlaAlaHisLeu 926
      :::::::::::::::::::::
Db 2965 GGGCAAGTCAGTTCATTGCTCTCTGACTATGCCAAAGCCAAATATCAGCAGCCACATC 3024
      :::::::::::::::::::::
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
      :::::::::::::::::::::
Db 3025 ATCATGATCATTTGAAAAAACCCCTTTGATTGACAGCTACAGCAGGAAGCGCTAATGCCG 3084
      :::::::::::::::::::::
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyRProCysArgPro 966
      :::::::::::::::::::::
Db 3085 AACACATTGGAAGAAATGTACATTTTGGTGAAGTTGATTCACTATCCACCCGACCG 3144
      :::::::::::::::::::::
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
      :::::::::::::::::::::
Db 3145 GACATCCAGTCTTCAGGACTCAGCTGGAGTGAGAGGCCAGACGCTGGCTCTG 3204
      :::::::::::::::::::::
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyRAsp 1006
      :::::::::::::::::::::
Db 3205 GTGGGCGACAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGAC 3264
      :::::::::::::::::::::
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
      :::::::::::::::::::::
Db 3265 CCCTTGGCAGGGAAGTGTCTTGTATGGCAAGAAATAAAGCGACTGAATGTTTCAGTGG 3324
      :::::::::::::::::::::
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
      :::::::::::::::::::::
Db 3325 CTCGAGCACACCTGGGCATCTGTTCCAGAGGACCCATCTCTGTTGACTCGAGCATTTGCT 3384
      :::::::::::::::::::::
Qy 1047 GluAsnIleAlatyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
      :::::::::::::::::::::
```

```
Db 3385 GAGAACATTCCTCTATGGACACAACAGCCGGTGGTGTCTACAGGAAGAGATCGTGAGGGCA 3444
      :::::::::::::::::::::
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyRAsnThrGln 1086
      :::::::::::::::::::::
Db 3445 GCAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCCCTAATAAATATAGCACTAAA 3504
      :::::::::::::::::::::
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
      :::::::::::::::::::::
Db 3505 GTAGGAGACAAAAGAACTCAGCTCTCTGCTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564
      :::::::::::::::::::::
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
      :::::::::::::::::::::
Db 3565 GCCTTTGTAGACGCTCATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624
      :::::::::::::::::::::
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
      :::::::::::::::::::::
Db 3625 GAAAGTGAAGAGTTGTCGAAGAGCCCTGGACAAAGCCAGAGAGCCGCGACCTGCATT 3684
      :::::::::::::::::::::
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleValValLeuHisAsn 1166
      :::::::::::::::::::::
Db 3685 GTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTGGTGTTCAGAAAT 3744
      :::::::::::::::::::::
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyRphe 1186
      :::::::::::::::::::::
Db 3745 GCAGAGTCAAGAGAGCATGGCAGCATCAGCAGCTGCTGGCAGAAAGGCATCTATTTT 3804
      :::::::::::::::::::::
Qy 1187 LysLeuValAsnAlaGlnSer 1193
      :::::::::::::::::::::
Db 3805 TCATGGTCAGTGCAGGCT 3825
      :::::::::::::::::::::
RESULT 6
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:
Pred. No.: 2,65e-310 Length: 4317
Score: 3549.00 Matches: 680
Percent Similarity: 74.65% Conservative: 239
Best Local Similarity: 55.24% Mismatches: 270
Query Match: 59.03% Indels: 42
DB: 9 Gaps: 8

US-09-873-409-6 (1-1195) x US-10-044-671-1 (1-4317)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
      :::::::::::::::::::::
Db 217 ATGTTGGTGGGACATGGCTGCCATCCATCCATGAGTGCATCCCTCTCATGATGCTG 276
      :::::::::::::::::::::
Qy 21 ValLeuGlyGluMetSerAsnLeuIleSerGlyCysLeu----- 34
```

Db 277 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCAAGGAATTTCAAGAAACAAAACCTTTT 336  
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 337 CCAGTTATTAATTAAGAAAGTATTAGCAACATATACACACATTTTCATCAACCATCTGGAG 396  
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
Db 397 GAGGAAATGACCACGATATCCCTATTATTACAGTGGGATCGGTGCTGGCTGGCT 456  
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrIlysergileArg 79  
Db 457 GCTTACATCCAGGTTTCATCTCGTGGCTGGCAGCAGGAAGACAGATCTCAAAATTTAGA 516  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
Db 517 AAACAATTTTTTCATGCTATCATGCGACAGAGGATGGCTGTTCAGCTGCATGACGTT 576  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 577 GGGAGCTTAACACCCGGCTCACAGACGATGCTCCAAAATCAATGAAGGAATTTGGCGAC 636  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 637 AAAGTTGGAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696  
Qy 139 ValLysGlyTyrIleThrLeuValThrLeuSerThrSerProIleuMetAlaSer 158  
Db 697 ACACCTGGTTGGAAGTAAACCTTTGATTTTGGCCATCATGCCCTGTTCTTGACATTTCA 756  
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 757 GCGCCATCTGGCAAGATACTATCTTCTTACTGATTAAGAATCTCTGGCCATGCA 816  
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 817 AAAGCTGGAGCAGTAGCTGAAGAAGCTTATGACAGCAATCAGAACTGTGATGCTTTGGA 876  
Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
Db 877 GGACAAAGAAAGAACTTTGAAAGGTACACAAAAATTTAGAAGAGCTAAAGAAATTTGGG 936  
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238  
Db 937 ATAAAGAAAGCTATCACGCCCAACATTTCTATGTTGGTCCGCTTCTTATTGATCTATGCA 996  
Qy 239 ThrTyrGlyLeuAlaPheTyrIleGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258  
Db 997 TCATATGCTCTGGCTTCTGGTATGGGACCTCTCTGGTCTCTCCAGTGAA-----TAT 1050  
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278  
Db 1051 ACTATTGGACAGTACTCACTGCTCTTCTTCTGTTATTAATTTGGGGCTTTTAGTATTGA 1110  
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298  
Db 1111 CAGGCATCCCAAGCATTGAGCAATTTGCAACGCAAGAGGACGACCTTATGAAATCTTC 1170  
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
Db 1171 AAGATAATTTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACAGAT 1230  
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
Db 1231 AATATTAAAGGAAATTTGGAATTTCAAAATATGTTTCACTTCAGTTTACCCTTCTCGAAAAGAA 1290  
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358  
Db 1291 GTTAAGATCTTAAGGGTCTCAACCTGAAGTTCAGAGTGGCAGAGTGGCGCTGGTT 1350  
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspPro 378

Db 1351 GGGAAACAGTGGCTGGGGAAGACGACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1410  
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398  
Db 1411 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCAATAATGTAAGGCATCTT 1470  
Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418  
Db 1471 CGGGAATTTACTGGTGTGGTGGTGGAGGCGCTGTGTTTGGCCACCACGATAGCTGAA 1530  
Qy 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438  
Db 1531 AACATTGCTATGGCGCGCAAAATGTCCATGGATGAGATTGAGAAAGCTGTTAAGGA 1590  
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458  
Db 1591 GCCAATGCTATGATTTTATCATGAACCTACCTAATAATTTGACACTCTGTTGGAGAG 1650  
Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478  
Db 1651 AGAGGGGCGCGCTGAGTGGTGACAGAAACAGAGAATCGCCATTTGCTCGGCGCTGGTT 1710  
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498  
Db 1711 CGCAACCCCAAGATTTCTCTGCTGATGAGGCAACGTCACTCTGGACACTGAAAGTGAA 1770  
Qy 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518  
Db 1771 GCAGTGGTTTCAGTGGCTGGATGAAGCCAGCAAAAGCCGAGCTACCATTTGTAGTGT 1830  
Qy 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538  
Db 1831 CATCGTTGCTACAGTTCGTAAATGCCGATGTCAITGCTGTTTGTATGATGAGTCAIT 1890  
Qy 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558  
Db 1891 GTGGAGAAAGAAATCATGATCAATCATGAAGAGGAGGAGGATTTTACCTCAAACTGTC 1950  
Qy 559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578  
Db 1951 ACAATGCCAG---ACAAGAGGAATGNAATTGATTAGAAATGCCACTGGTGAATCCAA 2007  
Qy 579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597  
Db 2008 AGTGAAGTGTATGCTTGGAAATG---TCTCCAAAAGATTCAAGGTCAGGTTTAATAAAA 2064  
Qy 598 -----AspLysAlaGluGluSer 603  
Db 2065 AGAAGATCAACTCGCAGGAGTATACATGCCACCAAGGCCAAGCAAGCTTGGTGACA 2124  
Qy 604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 623  
Db 2125 AAAGAGGACTTGAATGAGAAATGTTCTTTCAGTTTCTCTGGAGGATTTCTGAAGCTGAAC 2184  
Qy 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643  
Db 2185 TCAACTGAATGGCTTATTTTGTGGTGGTATATTTGTGCTATTATATAAACCGAGGCTG 2244  
Qy 644 HisProValPheSerIleIlePheAlaLysIleIleIleThrMetPhe---GlyAsnAsnAsp 662  
Db 2245 CAACACGACATTTTCAATAATAATTTTCAAGGATTTATAGGATCTTTTACCCGAGATGAGGAT 2304  
Qy 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682  
Db 2305 CCTGAAACAAAACGACAGAAATGATTAACATGTTTCTGTTATGTTTCTAGTCTTGGAAAT 2364  
Qy 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702  
Db 2365 ATTTCTTTTATTACATTTTCTCCAGGCTTCACATTTGGCAAGCTGGGGAGATCCTC 2424  
Qy 703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 722  
Db 2425 ACTAAGCGGCTTCGATACATACATGTTTTCAGATCCATCTCTGAGACAGGATGTACGCTGTT 2484





Db 331 ATGACCAAGATATGCTATTATTTACAGTGGAAATTTGGTGTGGGGTGTGCTGCTGCTTAC 390  
Qy 62 ILeGlnIleSerLeuTriPleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 391 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAACAAATACACAAAATTTAGAAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGCTATTAATGCGACAGAGATAGGCTGTTTGTATGTCACCATGTTGGGGAG 510  
Qy 102 LeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 511 CTTAACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGCACAAAT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGTTCTTTCAGTCATGCAACATTTTTCACCTGGGTTTATAGTAGGATTACAGCT 630  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 160  
Db 631 GGTGGAGCTAACCTTGATTTGGCCATCAGTCCTGTTCTTGGACTGTGCTGCTGCT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaLysSerLysAla 180  
Db 691 GTCTGGGCAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGACCTGTGATTCATTTGGAGACAA 810  
Qy 201 GluLysGluLeuGlnArgPyrThrGlnAenLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 811 AAGAAGAACTTTGAAAGGTACAAACAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAG 870  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValLysPhePheMetAsnGlyThrLys 240  
Db 871 AAAGCTATTACAGCCAAATATTTCTATAGTGTGCTGCTTCTCTGCTGATCATGCTTAT 930  
Qy 241 GlyLeuAlaPheTrpPyrGlyThrSerLeuIleLeuAenGlyGluProGlyTrpThrIle 260  
Db 931 GCTCTGGCCCTTCTGGTATGGACCACTTGGTCCCTCTCAGGGAA-----TATTCAT 984  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerPyrCysIleGlyAlaAla 280  
Db 985 GCACAAGTACTCACTGTAATCTTTCTGATTAAATTTGGGGCTTTTAGTCTGCGACAGCA 1044  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1045 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104  
Qy 301 IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTrpLysProGluSerIle 320  
Db 1105 ATTGATAATAAGCAAGTATTACAGCTATTGCAAGGTGGGCACCAACACAGATAATATT 1164  
Qy 321 GluGlyThrValGluPheLysAenValSerPheAsnTrpProSerArgProSerIleLys 340  
Db 1165 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTACCATCTCGAAAGAAAGTTAAG 1224  
Qy 341 IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1225 ATCTTGAAGGGCTGAACCTGAAGGTGCGAGGTGGGCAGACGGTGGCCCTGGTTGGAAC 1284  
Qy 361 AenGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTrpAspProAspAsp 380  
Db 1285 AGTGGCTGTGGGAAGAGCAACAGCTGCAGCTGATGACAGGGCTCTATGACCCCAAGAG 1344  
Qy 381 GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisLysTrpArgAsp 400  
Db 1345 GGGATGTCAGTGTTCATGGACAGGATATTAGACCAATAATGTAAGTTTCTCTCGGAA 1404  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAenAsnIle 420  
Db 1405 ATCATTTGGTGTGGTGTGAGTCAGGAACCTGATTGTTGTCACACAGCATAGTGAACAT 1464

Qy 421 LysTrpGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAen 440  
Db 1465 CGCTATGCGCGTGAATGTCCACATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAT 1524  
Qy 441 AlaTrpAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly 460  
Db 1525 GCCTATGACTTTATCATGAACTGCTCATTAATTTGACACCTGTTGGAGAGAGAGGG 1584  
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1585 GCCCAGTTGAGTGGGCGAGAGCAGAGGATCGCATTCACAGTGCCTGTTGCGAAC 1644  
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1645 CCCAAGATCTCTCTGATGAGGCCACGTGAGCTTGGACACAGAAAGCGAGCAGTG 1704  
Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
Db 1705 GTTCAGGTGGCTCTGTAAGGCCAGAAAGGTGCGACCATTTGTGATAGTCATCGT 1764  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 1765 TTGCTCAGCTTCGTAATGCTGACGTCATCGTCTGCTTCGATGATGGAGTCATTTGGAG 1824  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTrpSerLysValMetSer 560  
Db 1825 AAGGAAATCATGATGAATCATGAAAGAGAAAGGCACTTACTTCAAACTTGTCACAATG 1884  
Qy 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTTAGAAATGACGCTGATGAAATGCCAAAGTGAAT 1944  
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
Db 1945 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTTAATAAGAAAAGATCAACT 2004  
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607  
Db 2005 CGTAGGAGTGTCCGTGGATCACAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
Qy 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGluTrp 627  
Db 2065 GATGAAAGTATACCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATG 2124  
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProValPhe 647  
Db 2125 CTTTATTTTGTGTGGTGTATTTTGTGCCATTATAATGGAGGCTGCAACAGCATTT 2184  
Qy 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 2185 GCATAATATTTTCAAGATTTATAGGGTTTTACAAAGAAATTTGATGATCCTGAAACAAA 2244  
Qy 667 LysHisAspAlaGluIleTrpSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
Db 2245 CGACAGAAATAGTAACCTGTTTTCACCTATTGTTTCTAGCCCTTGGAAATTTCTTTAT 2304  
Qy 687 SerTrpPheMetGluGlyLeuPheTrpGlyArgAlaGlyGluIleLeuThrMetArgLeu 706  
Db 2305 ACATTTTCTTCAAGGTTTCATTTTGGCAAGCTTGGAGAGATCTTCAACAGCGGCTC 2364  
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTrpGlnAspIleAlaTrpPheAspGluLysGlu 726  
Db 2365 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGTTTTCATGCCCTAAA 2424  
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
Db 2425 AACACCATCTGGAGCATTCGACTACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT 2484  
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAenAlaThrAsnMetGlyLeuSerValIle 766  
Db 2485 ATAGGTTCCAGGCTTGTGTAATTTACCCAGATATAGCAATCTTTGGGACAGGAATAT 2544





US-09-873-409-6 (1-1195) x US-09-306-417-1 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 1370 ATGGTGGTGGGAACITTTGGCTGCCATCCATGGGCTGGAGCTTCTCTCATGATGCTG 1429  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 1430 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTCA 1489  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 1490 AACATCACTAATAGATGATATCATGATATACAGGCTTTCTCATGAATCTCGAGGAAGAC 1549  
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 1550 ATGACCAGATATCCCTATTATTACAGTGGAAATTTGGTCTGGGGTGTGCTTCTGCTTAC 1609  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 1610 ATTCAAGTTTCAATTTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAAATTAGAAAACAG 1669  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 1670 TTTTTCATGCTATATAATGCGACAGGAGATAGGCTGGTTTGTGATGTCACCATGTTGGGGAG 1729  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 1730 CTTAACCCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTGTTGACAAAAT 1789  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 1790 GGAATGCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 1850 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGCTGCT 1909  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 1910 GTCTGGGCAAGATACTATCTTACTATGATAAAGAACTCTAGCGTATGCAAAAGCT 1969  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 1970 GGAGCAGTAGCTGAAGAGCTTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 2029  
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
DB 2030 AAGAAGAACTTGAAGAGGTACACAAAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAG 2089  
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
DB 2090 AAGGCTATTACAGCCAAATTTCTATAGTGTCTGCTTCTGCTGATCTATGCACTTAT 2149  
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
DB 2150 GCTCTGCGCTCTGCTGATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 2203  
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
DB 2204 GGACAAAGTACTCACTGATTTCTTTCTGTTAATTTGGGGCTTTTAGTGTGGACAGGCA 2263  
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
DB 2264 TCTCCAAGCATTTGAAGCATTTTGAANAATGAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323  
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
DB 2324 ATTGATAATAAGCAAGTATTGACACTATTGCAAGAGTGGGCACAAACACAGATAATATT 2383  
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
DB 2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTATCCCATCTCGAAAGAAAGTTAAG 2443

QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
DB 2444 ATCTTGAAGGGCTGAACTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAC 2503  
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
DB 2504 AGTGCTGTGGGAAGAGCACACACGTGCAGCTGCAGAGGCTCTATACCCCCACAGAG 2563  
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
DB 2564 GGGATGGTCACTGTTGATCGACAGATATTAGGACCATAAATGTAAAGTTTCTACGGAA 2623  
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420  
DB 2624 ATCAITGGTGTGGTGCAGGAACCTGTATTGTTTGCACCACCATAGCTGTAACACATT 2683  
QY 421 LysTyrGlyArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440  
DB 2684 CGCTATGCCCGTGAATAATGTCCACATGATGAGATTGAAAGCTGTCAAGAGGCAAT 2743  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460  
DB 2744 GCCTATGACTTTTATCATGAACTGCTCTATAAATTTGACACCTGTTGCGAGAGAGGG 2803  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
DB 2804 GCCCAGTTGAGTGGTGGGAGAGCAGAGGATCGCAATTCGACGTGCTGTTGTCGAAC 2863  
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
DB 2864 CCAAGATCCTCTGCTGATGAGGCCACGTGAGCTTGGACACAGAAAGCGAAGCATG 2923  
QY 501 ValGlnAlaAlaGluLysAlaSerLysGlyArgThrThrIleValAlaAlaHisArg 520  
DB 2924 GTTCAGTGGCTCTCGATAAGGCCAGAAAGGTGCGACCACCATTTGTGATGCTCATGT 2983  
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
DB 2984 TTGCTACAGTTTCGTAATGCTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3043  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
DB 3044 AAGGAAATCATGATGAATCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCAATG 3103  
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
DB 3104 CAGACAGCAGGAAATGAAAGTTGAATTTAGAAATGCGAGCTGATGAATCCAAAGTGAAT 3163  
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
DB 3164 GATGCTTGGAAATGCTTCAATGATTCACAGATCCAGTCTAATAGAAAGAAAGTCAACT 3223  
QY 590 ----LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607  
DB 3224 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTG 3283  
QY 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627  
DB 3284 GATGAAGTATACCTCCAGTTCCTTTGGAGGATTTAGAGCTTAAATTTAACTGAATGG 3343  
QY 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
DB 3344 CTTATTGTTGTTGGTGTATTGTCATATAAATGGAGGCTGCAACAGCATTT 3403  
QY 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
DB 3404 GCAATAATATTTTCAAGATTTATAGGGTTTTTACAAGAAATTTGATGATCCTGAAACAAA 3463  
QY 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
DB 3464 CGACAGAATAGTAACCTGTTTCTCACTATTGTTTCTAGCCCTTGGAAATTTATTTCTTTAT 3523

```
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluLeuThrMetArgLeu 706
Db 3524 ACATTTTCTTCAAGGTTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGCTC 3583
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db 3584 CGATACATGGTTTCCGATCCATGCTCAGACAGAGTGTGGTTGGTTGATGACCCCTAA 3643
QY 727 AspSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db 3644 AACACACTGGAGCATGACTACAGGCTCGCAATGATGCTCTCAAGTTAAAGGGGCT 3703
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766
Db 3704 ATAGGTTCCAGGCTTGTGTAATATCCCAATATAGCAATCTTGGGACAGGAATAAT 3763
QY 767 IleSerPheIleTyrGlyTrpGluMetThrPheIleLeuSerIleAlaProValLeu 786
Db 3764 ATATCTTCTATGTTGGCAACTAACACTGTTACTCTTAGCAATTGACCCATCAT 3823
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaLeuLysAspLysGln 806
Db 3824 GCATAGCAGGAGTGTGTAATATGAATAATGTTCTGGCAACGACTGAAGAATAGAAA 3883
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db 3884 GAACTAGAGGTTCTGGGAAGATCGCTACTTGAAGSAATAGAAAACTTCCGAAACGTTGTT 3943
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 846
Db 3944 TCTTTGACTCAGGAGCAGAAATTTGAACATATGTAATGTCTGACAGATTTGACAGTACCATAC 4003
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db 4004 AGAACTCTTTGAGGAAGCAGACATCTTTGGAAATACATTTTCTTCACCCAGGCAATG 4063
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db 4064 ATGTAATTTTCTATGCTGATGTTTCCGGTTTGGAGCTTACTTTGGTGGCACATAAATC 4123
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db 4124 ATGAGCTTTGAGGATGTTCTGTAGTATTTTACGTGTTCTGTTGGTGCATGGCGGTG 4183
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db 4184 GGGCAAGTCACTTATTTGCTCTGACTATGCCAAGCCAAATATCAGCAGGCCCATC 4243
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db 4244 ATCATGATCATTTGAATAAACCCCTTTGATTGACAGCTACAGCAGGAAGCCCTAATGCGG 4303
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db 4304 AACACATTGGAAGGAATGTACATTTTGGTGAAGTGTATTCAACTATCCACCCGACCG 4363
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 4364 GATCCCATGCTTTCAGGAGCTGAGCTGGAGGTGAAGAGGGCCAGAGCTGGGCTCTG 4423
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 4424 GTGGCAGCAGTGGCTGTGGGAAGAGACACAGTGTCCAGCTCTCTGGAGCGGTTCTACGAC 4483
QY 1007 ProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db 4484 CCCTTGGCAGGGAAGTGTCTGTATGTGGCAAGAAATTAAGCGACTGAATGTTCACTGG 4543
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 4544 CTCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTGTTTGACTGCAGCATGCT 4603
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
```

```
Db 4604 GAGAACATTGCTATGGAGACAAACAGCCGGTGGTGTCTACAGGAAGAGATCGTGGGGCA 4663
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 4664 GCAAGGAGGCCAACACATACATGCTTCTATGAGTCACTGCTCTAATAAATATAGCACTAA 4723
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 4724 GTAGGAGACAAAGGAATCAGCTCTCTGTGTGGCCAGAAACAACGCATTGCCATAGCTCGT 4783
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 4784 GGCCTTGTGTAGACAGCCTCATATTTTGTCTTTGGATGAAGCCACGCTCCTGGATACA 4843
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 4844 GAAAGTGAAGAAGTGTCTCCAAAGCCCTGGCAAGCCAGCAAGAGCCGACCTGCAT 4903
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 4904 GTGATTGCTCACCGCTGTCCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGAAAT 4963
QY 1167 GlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 4964 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5023
QY 1187 LysLeuValAsnAlaGlnSer 1193
Db 5024 TCAATGGTCAGTGTCCAGGCT 5044

RESULT 9
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta91msAI
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: mSAI mdrl cdna
; FEATURE:
; NAME/KEY: misc_feature
```

; LOCATION: (5215)..(5774)  
; OTHER INFORMATION: 3'-LTR  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (5775)..(8630)  
; OTHER INFORMATION: plasmid backbone (pUC)  
US-09-306-417-2

## Alignment Scores:

Pred. No.: 1,67e-309 Length: 8630  
Score: 3545.00 Matches: 681  
Percent Similarity: 75.55% Conservatives: 246  
Best Local Similarity: 55.50% Mismatches: 264  
Query Match: 58.97% Indels: 36  
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-306-417-2 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 1370 ATGGTGGTGGAACTTTGGTGGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 1430 GTTTGGAGAAATGACAGATATCTTTGCAATGCGAGAAATTTAGAGATCTGATGTCA 1489  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 1490 AACATCACTAATAGAGTATGATATCAATGATACAGGGTTCCTTCATGAATCTGGAGGAAGAC 1549  
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 1550 ATGACCAAGGTATGCCATTATTACAGTGGAAATTTGGTGGGGTGGTGGTCTTAC 1609  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgIleGlyGln 81  
DB 1610 ATTCAGGTTTCATTTGGTGGCCCTGGCAGCTGGAGAACATAACAAAAATTAGAAACAG 1669  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 1670 TTTTTCATGCTATATATGACACAGGAGATAGGCTGTTGATGTGCACGATGTGGGGAG 1729  
QY 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspIle 120  
DB 1730 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAATGAAGCTTATTGGTGACAAAAAT 1789  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 1790 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 1850 GGTGGAGACTAACCTTGATTTGGCCATCAGTCTGCTTCTTGCACTGTCAGCTGCT 1909  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerLafatyrSerLysAla 180  
DB 1910 GTCTGGGCAAGACTACTATCTTCAATCTGATATAAGAACTCTTAGCGTATGCAAAAGCT 1969  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 1970 GGAGCAGTAGTCTGAAGAGGTCCTGGCAGCAATTTAGAACTGTGTGATTTGGAGACAA 2029  
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
DB 2030 AAGAAGAAGCTTGAAGGTHACACAAAAATTTAGAGAGCTAAAGAATTGGGATTAAG 2089  
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
DB 2090 AAAGCTATTACAGCCCAATTTCTATAGGTGGCTCTTCTGCTGATCTATGATCTTAT 2149  
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 260  
DB 2150 GCTCTGGCCTTCTGGTATGGGACCACTTGGTCTCTCTCAGGGAA-----TATTCTATT 2203

QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
DB 2204 GGACAAGTACTCACTGATTTCTTTCTGATTAAATGGGGCTTTTAGTGTGGCAGGCA 2263  
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
DB 2264 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATCAAAATCTTCAAGATA 2323  
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
DB 2324 ATTGATAATAAGCAAGTATTGACAGCTATTGAGAGTGGGCACAAACAGATAATATT 2383  
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
DB 2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443  
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyThrValAlaLeuValGlyLeu 360  
DB 2444 ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGCAGACGGTGGCCCTGGTTGAAAC 2503  
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
DB 2504 AGTGGCTGTGGAGAGGACACACAGCTCAGCTGATGCAGAGGCTCTATGCCCCACAGAG 2563  
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
DB 2564 GGGATGGTCAGTGTGATGGACAGGATATTAGGACCAATAATGTAAGGTTTCTACGGGAA 2623  
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
DB 2624 ATCATTTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2683  
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
DB 2684 CGCTATGGCGGTGAAATGTCCACATGATGAGATTGAGAAAGCTCTCAAGGAGGCAAT 2743  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460  
DB 2744 GCCTATGACTTTATCATGAACCTGCCTCATATAATTTGACACCCCTGTTGGAGAGAGGG 2803  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
DB 2804 GCCCAGTTGAGTGGTGGGAGAGCAGAGGATCGCCATTTGCAGTCCCTCGTTCGCAAC 2863  
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
DB 2864 CCCAAGATCCTCTGCTGGATGAGCCAGCTGAGCTTGGACACAGAAAGCGAAGCAGATG 2923  
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
DB 2924 GTTCAGTGGCTCTGGATAAGGCCAGAAAAGTTCGACCCACCATTTGTGATGCTCATCGT 2983  
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
DB 2984 TTTGCTACACCTCGTAATGCTGACGCTCATCGTGGTTCGATGATGGAGCTATTGGGAG 3043  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
DB 3044 AAGGAAATCATGATGAATCATGAAGAGAAAGGATTTACTTCAAACTGTGCACAATG 3103  
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
DB 3104 CAGACAGCAGGAATAAGAACTTGAATTTAGAAAATGACAGTGAATCCAAAGTGAAT 3163  
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
DB 3164 GATGCTTGAAGTGTCTTCAAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3223  
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607  
DB 3224 CGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3283

Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp	627
Db	3284	GATGAAGTAGTACCTCCAGCTTCCTTTTGGAGGATATGAAGAGTTTAACTGAATGG	3343
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	3344	CCATTATTTGTGTGGTGATATTTGTGCCATATATAATGGAGGCTTGCACACGACATTT	3403
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAspLysThrThrLeu	666
Db	3404	GCAATATATTTTCAAGATTATAGGGGTTTTTACAGAAATTGATGATCCTGAAACAAA	3463
Qy	667	LysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	3464	CGACAGAAATAGTAACCTGTTTTTCACTATTGTTCTTAGCCCTTGGAAATTATTCTTTTATT	3523
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	3524	ACATTTTTCCTTCAAGGTTTCACATTTGCGAAGCTGGAGAGATCCTCACCACGCGCTC	3583
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu	726
Db	3584	CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCATAA	3643
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	3644	AACACCACTGGAGCATTTGACTACAGGCTCGGCCAATGATGTGCTCAAGTTTAAAGGGCT	3703
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	3704	ATAGTTTCCAGGCTGCTGTAATATCCAGAAATATAGCAAAATCTTGGGACAGGAATAATT	3763
Qy	767	IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	3764	ATATCCTTCATCATGGTTGGCACTAAACACTGTACTCTTAGCAATTTGTACCCATCATT	3823
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3824	GCAATAGCAGGAGTTGTTGAAATGAAATGTTGTCTGGACAGCACTGAAAGATAAGAAA	3883
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3884	GAACTAGAAGGTCCTGGAGAGATCGCTACTCGAAGCAATAGAAAACCTCCCGAACCGTGT	3943
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetGlnThrGlnHis	846
Db	3944	TCCTTTGACTCAGGACGAGAAGTTTGAACATATGATGCTCAGAGTTTCAGGTCACCATAC	4003
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	4004	AGAAACTCTTTGAGGAAGCACACATCTTTTGGAAATTACATTTTCTCTCACCCAGGCAATG	4063
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	4064	ATGTATATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGGCTACTTGTGTGGCACATAAACTC	4123
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	4124	ATGAGCTTGGAGGATGTTCTGTAGTATTTTTCAGCTGTGTTCTTTTGGTGCCATGCCGCTG	4183
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
Db	4184	GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAAATATATCAGACGCCACATC	4243
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro	946
Db	4244	ATCATGATCATTTGAAAACAAACCCCTTTGATTGACGCTACAGACGGAAGGCTTAATGCCG	4303
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966
Db	4304	AACACATTGGAAGAAATGTCACATTTGTTGAAGTTGTATTTCACTATCTCCACCCGACCG	4363
Qy	967	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyThrValAlaPhe	986

```

Db 4364 GACATCCAGCTGCTTCAGGAGCTGAGCCCTGGAGGTGAGAGAGGGCCAGACCGTGGCTCTG 4423
Qy 987 ValGlySerSerGlyCysGlyLysserThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 4424 GTGGGCAGCAGTGGCTGTGGGAAGACACAGAGTGTCCAGCTCCTGGAGCGGTCTTACGAC 4483
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db 4484 CCCTTGGCAGGAAAGTGCTGTGATGGCAAAAGATAAAGCGACCTCAATGTTTCAGTGG 4543
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 4544 CTCGAGCACACCTCGGCACTCGTGTCCAGAGGCCCATCCTGTTGACTCGCAGCATTCCT 4603
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db 4604 GAGAACATTCCTATGTGAGACACACCCGGGTGTGTACAGGAAGAGATCGTGAGGGCA 4663
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 4664 GCAAAGGAGGCCAACATACATGCTTCATCGAGTCACCTGAATAAATATAGCACTAAA 4723
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 4724 GTAGGACACAAAGGAACCTCAGCTCTCTGTGTGCCAGAAACAACGCAATTGCCATAGCTCGT 4783
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 4784 GCCCTGTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCGGATACA 4843
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 4844 GAAAGTGAAGAGGTGTGCCAAGAGCCCTGGACAAGCCAGAGAGCGCGCACCTGCAAT 4903
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsn 1166
Db 4904 GTGATTCTCCAGCCGCTGTCCACCATCCAGAAATGCAGACTTAAATAGTCGTGTTTCAGAA 4963
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuLeuAspAsnArgPhe 1186
Db 4964 GCGAGTGAAGAGCATGCGACCCATCAGAGCTGTGGCCACAGAAAGGCATCTATTTT 5023
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 5024 TCAATGTCAGTGTCCAGGCT 5044

RESULT 10
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCES: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA

```

ORGANISM: RATTUS RATTUS  
US-09-769-097-3

## Alignment Scores:

Pred. No.: 1,6e-309 Length: 4425  
Score: 3540.50 Matches: 680  
Percent Similarity: 75.53% Conservative: 243  
Best Local Similarity: 55.65% Mismatches: 266  
Query Match: 58.89% Indels: 33  
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x US-09-769-097-3 (1-4425)

```
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 499 ATGCTGCGGAACCTCTGGCGGCATTATCCATGGAATTCGCTCCCACTTATGATGCTG 558
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 559 GTCTTTGGAGACATGACAGATAGCTTTTGGCAAT-----GTAGGAACAAACCGTAGTAG 612
QY 41 SerPhePheArgLeuThr----- 46
DB 613 AGTTTCTCAATCTACAGACATATATGCCAAGCTGGAGGACGAATGACCAGTAGGCC 672
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 673 TACTATTACAGGGCATTTGGTGGCGGTGCTCATCGTTGCCCTACATCCAGGTTTCCACT 732
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
DB 733 TGGTGGCTGGCAGCTGGGACAAATACACAAGATTAGGCAGAAAGTTTTTCCATGCCATC 792
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 793 ATGAATCAGAGATAGCTGTTTGGCTGCTGATGAGCTGGGAGGCTCAACACCCGGCTC 852
QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
DB 853 ACAGATGACGTCTCCAAATTAATGAAGGAATGGTGACAAAATTTGGAATGCTTTTCAG 912
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
DB 913 GCAATGGCAACATTTTTTGGTGGTGTATTAATAGGAATTTACTCGCGGCTGGAAAGCTA 972
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaLacysSerArgMet 165
DB 973 CTGTGATTTGGCCATCAGCCCTGTTCTTGGACTGTGACTGTGATTTGGGCAAGATA 1032
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
DB 1033 TTGCTCTTCAATTTACTGATAAGGAACCTCCAGGCTTATGCAAAAGCTGGAGCAGTTG 1092
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
DB 1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTTGCCCTTTGGAGGACAAAAGAAAGAACT 1152
QY 206 ArgTyrThrClnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
DB 1153 AGGTACAATTAACAAATTTGGAAGAGCTTAAAGGCTTGGGATAAAGAAAGCTATCAGGC 1212
QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTip 245
DB 1213 AACATTTCCATGGGTGAGCTTTTCTGCTTATCTATGTCATCATATGCTCTGGCATCTG 1272
QY 246 TyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
DB 1273 TATGGGACTTCTTGGTCTATCTCAAAAGAA-----TACACTATTGACAAGTGTCTCACT 1326
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
DB 1327 GTCTTTTCTGTATTAATTTGGAGCATTCAGTGTGGGCGAGCATCTCCAAATATTGAA 1386
```

```
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysIysPro 305
DB 1387 GCCTTTCGCCAATGCTTAGAGGACGCTTATGAAGTCTTATAGTATAATTAAGATAGCCC 1446
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
DB 1447 AGTATAGACAGCTTCTCAAGAGTGGGCAAAACCCGCAACATACAAAGAAATTTGGAA 1506
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
DB 1507 TTCAAAAATATTCACTTCACTTACCGTCTCGAAAAGACGTTTCAGATCTTGAAGGCCCTC 1566
QY 346 AsnLeuArgIleLysSerGlyGlnThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
DB 1567 AACCTGAAGGTGAAGAGCGGCGAGCGTGGTGGTGGCAACAGTGGCTGTGGGAAA 1626
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
DB 1627 AGCACAACTGTTCAGCTGCTGAGAGGCTCTACAGCCCATAGAGGCGGAGGTCAGTATC 1686
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
DB 1687 GACGGACAGACATCAGGACCATCAATGTGAGGTATCTCGGGGAAATCATTTGGGTGGTG 1746
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
DB 1747 AGTCAGAAACCGGTGCTGTTGCCACCAATTTGCCGAAAACATTCGCTATGCGCCGAGAA 1806
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
DB 1807 AACGTCACTGATGATAGAGAAAGCTGTCAAGGAAGCCCAATGCCATGATTTTCATC 1866
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly 465
DB 1867 ATGAACCTCCCAACAAATTTGCACACCTGTTGGTGAGAGAGGCGCGAGTGGTGG 1926
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
DB 1927 GGACAGAAACAGAGGATCCCATTTGCCGGGCGCTGTCCGCAACCCCAAGATCTCTTTT 1986
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 505
DB 1987 TTGATGAGCGCACGTCAGCTTTGGACAGAAAGGAAGCCGTGGTTTCAGGCGGCTCTG 2046
QY 506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
DB 2047 GATAAGCTTAGAGAGCGCGGACCACTTGTGATAGTCTCACGCTTGTCTACAGTTCCG 2106
QY 526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
DB 2107 AATGCTGATCATTCCTGTTTGTGATGGTGTCTATTGTGGAGCAAGAAATCATGAT 2166
QY 546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln-----MetThrTyr 575
DB 2167 GAGCTCATGAGAGAAAGAAATTTACTTCAAACTTGTCTACATCAGACAGAGCAAGAAAT 2226
QY 562 AspIleLysAlaAspGluGlnMetGluSer-----MetThrTyr 575
DB 2227 GAATTTGAATTAGGAATAAGCTTGTGAATCTAAAGATGGAATGATATGTTGGACATG 2286
QY 576 SerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIleLys 593
DB 2287 TCTTCAAAAGATTTCRGATCCAGTCTTAATAAGAAAGATCAACTCGCAAAAGATCCCGT 2346
QY 594 SerAspPheIleAspLysAlaGluSerThr---GlnSerLysGluLysSerLeuPro 612
DB 2347 GGGCCACATGATCAAGACGGGGAACCTTAGCAAAAGAGGCTCTGGATGACGCGTACCT 2406
QY 613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeu 632
DB 2407 CCAGCTTCTTTTGGCGGATCTGAAATTTCACTGAATGGCTTATTTTGGTGGTT 2466
QY 633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
```



```
Db 2467 GGTGTTATTTGTCCTCAATAAATGAGGCTTGCACACGACCATCTCCATAATATTTC 2526
QY LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAlaGlu 671
Db 2527 AAGGTTGTAGGGTTTTTACAAAATAATGACACCCCTGAAATCCAGCGGAGAACACAC 2586
QY 672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
Db 2587 TTGTTTCTTTATGTTCTGATCCCTTGGATCATCTCTTTTATTAGTTTTCCTTCAG 2646
QY 692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
Db 2647 GGCTTCACATTTGGCAAGCTGAGAGATCCTCACCAAGGCACTCCGATACATGCTCTTC 2706
QY 712 LysAlaMetLeuTyrGlnAspIleAlaTyrPheAspCyluLysGluAsnSerThrGlyGly 731
Db 2707 AAATCCATGCTGAGACAGGACATAAGCTGTTGATGACCTAAATAACACACAGAGCG 2766
QY 732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
Db 2767 CTGACACACAGGCTTGCCATGACGCTGCTCAAGTGAAGGGGCTACAGGCTCTAGGCTT 2826
QY 752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
Db 2827 GCTGTTATTACCCAGAACATAGCAATCTTGGACAGGACATCATCATCCCTGATCTAC 2886
QY 772 GlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
Db 2887 GGCTGGCAATTTGACACTTTTACTCTAGCAATTTGTTCCCATCATTTGATAGCAGGAGT 2946
QY 792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
Db 2947 GTTGAAATGAAATGTTGCTGCAACAGCCCTGAAGATGAAGAGGAACTAGAAGGTTCT 3006
QY 812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
Db 3007 GCGAAGATCGCTACAGAAGCAATTTGAAACTTTTCGCACCTCGTCTCTTTGACTCGGAG 3066
QY 832 LysAlaPheGluGlnMetTyrGluLysMetLeuGlnThrGlnHisArgAsnThrSerLys 851
Db 3067 CAGAAGTTTGAAACTATGATATGCCCAGAGCTGCGAGATACCATACAGAAATGCTTGAAG 3126
QY 852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
Db 3127 AAGCGCACGCTTTGGGATCATTCTCTTCCACAGGCCATGATGATTTCTCTCTAT 3186
QY 872 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 891
Db 3187 GCTGCTTTGTTCCGGTTTGATGCTACTTGTGGCACGAGAACTCATGACATTTGAAAT 3246
QY 892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
Db 3247 GTTCTGTAGTATTCTCAGCTATTGCTTGTGGTGCATGCGAGTGGGGAGGTCAAGTTCA 3306
QY 912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGlu 931
Db 3307 TTGCTCTGACTACGCGAAAGCCAAAGTCTCGCATCTCCCATCATCATCAGATCATTGAG 3366
QY 932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
Db 3367 AAATCCCTGAGATTGACAGTACAGACAGGAGGCTTGAAGCCTAATATGTTGGAAGGA 3426
QY 952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
Db 3427 AATGTGAAATTTAATGAGTTCATGTTCAACTATCCCAACCCGACCCCAACATCCCGTCTT 3486
QY 972 ArgGlyLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
Db 3487 CAGGGGCTGAGCCTAGAGGTGAGAAAGGCGACGCTGCGCCCTCGTGGGACAGGTGGC 3546
QY 992 CysGlyLysSerThrSerValGlnLeuGlnAlaGluLeuTyrAspProValGlnGlyGln 1011
```

```
Db 3547 TCGGGGAAGAGTACAGTGGTCCAGCTGCTTGGACGCGCTTCTATGACCCCAATGCGGAAACA 3606
QY 1012 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIle 1031
Db 3607 GGTGTTCTAGATGGCAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCACCTG 3666
QY 1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
Db 3667 GGCAATTGTCTCCAGAGGCCCATCTGTTTGTAGCTGCAGCATCGCCGAGAACATTCCTAC 3726
QY 1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1071
Db 3727 GGAGACAAACAGCCGCTGCTGCTCTCATAGAGAGATCGTGAAGGCAGCCCAAGGAGGCAAC 3786
QY 1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
Db 3787 ATCCACCACTTCATGACTCATCTGCTGAGAAATACACACAGAGTGGAGNCAAGGG 3846
QY 1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLys 1111
Db 3847 ACTCAGCTGTCGGGGGCGAGAGCAGCGCATCGCCATCGCGCGCCCTCGTCAGACAG 3906
QY 1112 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
Db 3907 CCTCACATCTTACTTCTGGATGAAGCGACATCATGCTCTGGATACGAGAGTGAAGAGTGC 3966
QY 1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
Db 3967 GTCCAGAAAGCGCTGCAAAAGCCAGGAGGAGCGCACCTGATCTGTATGTCGCGCACCGC 4026
QY 1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsnGlyLysIleLysGlu 1171
Db 4027 CTGTCCACCATCCAGAACCGACACTTGATCGTGGTGATTCAGAACCGCCAGTCAAGGAG 4086
QY 1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db 4087 CACGGCACCCACAGCAGCTGCTGGCCCAAGAGCATCTATTTCTCGATGGTCAAGTGTG 4146
QY 1192 GlnSer 1193
Db 4147 CAGGCT 4152

RESULT 11
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1
Alignment Scores: 2.39e-309 Length: 4369
Pred. No.:
```



Db	1447	AGTATAGACAGCTTCTCAAGAGAGTGGGCACAAACCCGCAACACATACAAGGAAATTTGGAA	1506
Qy	326	PhelyAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysIleLeuLysGlyLeu	345
Db	1507	TTCAAAATATTACCTTCAGTTACCGTCTCGAAGAAGAGCTTCAGATCTTGAAGGCGCTC	1566
Qy	346	AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys	365
Db	1567	AACCTGAAGGTGAAGCGGCGAGACGGTAGCCCTGGTGGCAACATTCGGTGGGAAA	1626
Qy	366	SerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal	385
Db	1627	AGCAAACTCTGCAGCTGTGACAGAGCTCTACGACCCCATAGAGGGCGAGGTCAGTATC	1686
Qy	386	AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal	405
Db	1687	GACGACAGGACATCAGGACCATCATGTGAGGTATCTCGCGGAAATCATTTGGGTGGT	1746
Qy	406	SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp	425
Db	1747	AGTCAGAAACCCGTGTGTGTCACCAAAATTCGCGGAAAAATTCGTATGGCCGAGAA	1806
Qy	426	AspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle	445
Db	1807	AACGTCAACATGATGAGATAGAGAAGCTGTCAAGGAAGCCAAATGCCTATGATTTCATC	1866
Qy	446	MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly	465
Db	1867	ATGAACTGCCCAAAATTTGACACCTGGTGGTGAGAGAGGGCGGACGTGAGTGGG	1926
Qy	466	GlyGlnSerGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle	485
Db	1927	GGACAGAAACAGAGGATCGCATTCGCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTG	1986
Qy	486	LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu	505
Db	1987	TTGGATAGCGCCACGTGACGCTTGGACACAGAAAGCGAGCGCTGGTTCAGCGCCCTCTG	2046
Qy	506	GluLysAlaSerLysGlyArgThrThrIleValAlaHisArgLeuSerThrIleArg	525
Db	2047	GATPAGGCTAGAGAGCCGACCAACATTTGTGTAGCTCACCGTTGTCTACAGTTGCG	2106
Qy	526	SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla	545
Db	2107	AATGCTGACGTCATTGCTGTTTGTATGGTGTGTCATTGTGTGAGCAAGGAAATCATGAT	2166
Qy	546	GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln	561
Db	2167	GAGCTCATGAGAGAGAAAGAAATTTACTTCAACATTGTCTATGACTCAGACAGCAGAAAT	2226
Qy	562	AspIleLysLysAlaAspGluGlnMetGluSer	575
Db	2227	GAAATGAATAGAAATGAGCTTGTGAATCTTAAAGCGGAATTCATAATGTGACATG	2286
Qy	576	SerThrGluArgLysThrAsnSerLeu	593
Db	2287	TCITCAAAGATTCCGGATCCAGTCTTAATAAGAAGAAGATCACTCGCAAAAGCATCGT	2346
Qy	594	SerAspPheIleAspLysAlaGluGluSerThr	612
Db	2347	GGGCGCATGATCAAGACGGGGAATTCAGCAAAAGAGGCTCTGGATGACGACGTACT	2406
Qy	613	GluValSerLeuLeuLysIleLysLeuAsnLysProGluTyrProPheValValLeu	632
Db	2407	CCAGCTTCCTTTTGGCGGATCCCTGAAGTTGAATTCACCTGAATAGCCCTATTTTGTGGT	2466
Qy	633	GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla	652
Db	2467	GGTGTATTTTGTCCATAATAATGAGGCTTGCAACACGACATCTCCATAATATTTTCA	2526
Qy	653	LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu	671

```
Db 2527 AAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAATCCAGCGCGCAGAACACAAAC 2586
Qy 672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
Db 2587 TTTGTTTCTTATTTGTCCTGATCCTTGGATCATCTCTTTCAATAGGTTTTTCTTCTCAG 2646
Qy 692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
Db 2647 GGCTTCACATTGGCAAAAGCTGGAGAGATCCTCACCACAGGAGCTCCGATACATGCTTTC 2706
Qy 712 LysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGly 731
Db 2707 AAATCCATGCTGACAGCAGCATAAGCTGTTGATGACCCCTAAAAACACCCAGCAGCG 2766
Qy 732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
Db 2767 CTGACCACCAAGGCTTGCCAAATGACGCTGCTCAAGTGAAGGGGCTACAGGGTCTAGGCTT 2826
Qy 752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
Db 2827 GCTGTTATTAACCAACATAGCAATCTTGGCAGGCGATCATATCCCTGATCTAC 2886
Qy 772 GlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
Db 2887 GGCTGGCAATTGACATTTTACTCTAGCAATTTGTTCCCATCATTTGCTATAGCAGGAGTG 2946
Qy 792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
Db 2947 GTTGAAATGAAATGTTGTCTGGCAAGCGCTGAAGATGAAGAGAACTAGAAAGGTTCT 3006
Qy 812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
Db 3007 GGGAGATTCGTACAGAGCAATTTGAACCTTTCGCACCTGCTCTCTTTGACTCGGGAG 3066
Qy 832 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 851
Db 3067 CAGAAGTTTGAACATATGATATGCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAG 3126
Qy 852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
Db 3127 AAGCGCAGCTGTTTGGGATCAGCTTTCTCCTTCACCCAGGCGCATGATGATTTCTCTAT 3186
Qy 872 AlaAlaGlyPheArgPheGlyAlaTyrIleLeuIleGlnAlaGlyArgMetThrProGluGly 891
Db 3187 GCTGCTGTTTCCGGTTTGATGCTTACTTGTGGCCAGCAACTCATGACATTTGAAAT 3246
Qy 892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
Db 3247 GTTCTGTTAGTATTCTCAGCTATTGTTCTTTGGTCCCATGGCAGTGGGCGCAGGTCAGTTC 3306
Qy 912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuGlu 931
Db 3307 TTCGCTCCTGACTACGGAAAGCCAAAGTCTCTGGCATGCCATCCCATCATCAGGATCATTTG 3366
Qy 932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
Db 3367 AAAATCCCTGAGTTGACAGTACAGCAAGCGAGGGCTTGAAGCCCTATATGTTGGAAGGA 3426
Qy 952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
Db 3427 AATGTGAAATTTAATGGAGTTCATGTTCAACTATCCACCCGACCCCAACATCCAGTGTCT 3486
Qy 972 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
Db 3487 CAGGGGCTGAGCCTAGAGGTGAAGAAAGGCGAGCGCTGGGCCCTCGTGGGCGAGGAGTGGC 3546
Qy 992 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGln 1011
Db 3547 TCGGGGAAGATACAGTGTCTCAGCTGCTTGGAGGCTTCTATGACCCCATGCCCGGAACA 3606
Qy 1012 ValLeuPheAspGlyValAlaAspAlaLysGlnLeuAsnValGlnTyrLeuArgSerGlnIle 1031
Db 3607 GTGTTTCTAGATGGCAAGAAATAAGCAACTCAATGTCCAGTGGCTCCGCGGCCCATCTG 3666
```

```
Qy 1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
Db 3667 GGCAATTGTCTCCAGAGCCCATCTGTTTGTAGCTGCAGCATCGCCGAGAACATTTGCTTAC 3726
Qy 1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1071
Db 3727 GGAGACAAACAGCGGTGCTGCTCTCATAGAGAGATCGTGAAGGCGACCAAGGAGGCCAAC 3786
Qy 1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
Db 3787 ATCCACCACTTTCATGACTCATCTGCTGAGAAATACAAACACAGATGGGAGACAAAGGG 3846
Qy 1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1111
Db 3847 ACTCAGCTGCTCGGGCGGAGAGCAGCGCATCGCCATCGCGCGCCCTCGTTCAGACAG 3906
Qy 1112 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
Db 3907 CCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTCGATACCGAGAGTGAAGGTC 3966
Qy 1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
Db 3967 GTCAGGAAGCGCTGGACAAAGCCAGGAGGAGCCGACCTGCTGATTTGATCGCGCACCGC 4026
Qy 1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1171
Db 4027 CTGTCCACCATCCAGAACCGACAGACTTGTGCTGTTGATTTCAGAACCGCCAGGTCAAGG 4086
Qy 1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db 4087 CACGGCACCCACCCAGCAGCTGCTGGCCCAAGAAAGGCATCTATTTCTCGATGGTCAAGT 4146
Qy 1192 GlnSer 1193
Db 4147 CAGGCT 4152
```

## RESULT 12

```
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2
```

```
Alignment Scores:
Pred. No.: 3,21e-309 Length: 4643
Score: 3537.50 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 263
Query Match: 58.84% Indels: 37
DB: 9 Gaps: 10
```

US-09-873-409-6 (1-1195) x US-10-072-621-2 (1-4643)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634
Qy 21 ValLeuGlyLeuMetSerAsp-----AsnLeuIleSer 31
```

[illegible]

1706	Db	AGTGGCTGTGGGAAGACAAACAGTCCACGTCCAGCTGTATGACAGAGCTCTATGACCCACACAGAG	1765
381	Qy	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
1766	Db	GGGATGCTCAGTCTTCATGCACAGGATATTAGCACCATATAAATGTAAGGTTTCTACGGGAA	1825
401	Qy	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	420
1826	Db	ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTTGGCCACACGATAGCTGAAACATT	1885
421	Qy	LysTyrClyArgAspValThrAspGluGluMetGluArgAlaAalaArgGluAlaAsn	440
1886	Db	CGCTATGGCCGTGAAATATGCACCTAGTAGATGTAGAAGAGCTGTCAAGGAAGCCAAAT	1945
441	Qy	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValClyGluLysGly	460
1946	Db	GCCTATGACTTTATCATGAAACCTGCCTCATATAATTTGACACCTGTTGGAGAGAGAGG	2005
461	Qy	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
2006	Db	GCCCAAGTTGAGTGGTGGCAGAACGAGAGATCGCCATTGCACGTGCCTCGTTGTCGAAC	2065
481	Qy	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
2066	Db	CCCAAGATCTCTCTGTGATGAGGCCAGCTCAGCCTTGGACACAGAAGCGAAGAGTG	2125
501	Qy	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
2126	Db	GTTTCAGGTGGCTCTGGATAAGGCCAGAAAGGTCGGACCCACCATTTGTATAGCTCATCGT	2185
521	Qy	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
2186	Db	TTGTCTACAGTTCTGTAATGCTCAGCTCATCGCTGGTGTTCGATGATGGAGTCATTGTGGAG	2245
541	Qy	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
2246	Db	AAAGGAAATCATGATGAACCTCATGAAGAGAAGGATTTACTTCAAACTTGTCACAATG	2305
561	Qy	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
2306	Db	CAGACACAGGAATGAGTTGNAATTAGAAAATGCAGCTGATGATGATCCAAAGTCGAATT	2365
575	Qy	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
2366	Db	GATGCCCTTGGAAATGCTTCCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT	2425
590	Qy	---LysSerIleLys---SerAspPheIleAspLysAlaGluLysSerThrGlnSerLys	607
2426	Db	CGTAGGAGTCTCGTGGATCAACAGCCCAAGACAGAAAGCTTTAGTACCAAGAGGCTCTG	2485
608	Qy	GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr	627
2486	Db	GATGAAAGTATACCTCCAGTTTCCTTTTCGAGATATAGAGCTAAATTTAACTGNAATGG	2545
628	Qy	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
2546	Db	CCTTATTTTGTGTGTGTATTGTCGCATTATATAATGAGGCGCTCGAACACGACATTT	2605
648	Qy	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
2606	Db	GCAATAATATTTTCAAAGATATATAGGGGTTTTTTTACAAAGAAATTTGATGATCCTGAAACAAA	2665
667	Qy	LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
2666	Db	CGACAGAAATAGTAACCTGTTTTCACCTATTGTTTCTAGCCCTTGGAAATATTATCTTTTATT	2725
687	Qy	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
2726	Db	ACATTTTTTCTCAGGGTTTTCACATTTTGGCAAAGCTGGAGATCCTCACCAAGCGGCTC	2785
707	Qy	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrrPheAspGluLysGlu	726
2786	Db	CGATACATGTTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGTATGACCTTAAA	2845

```
QY 727 AspSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 746
|||:|||||
Db 2846 AACACCACTGGAGCATGTACTACAGCTCCCAATGATGCTCTCAAGTTAAAGGGGCT 2905
|||:|||||
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766
|||:|||||
Db 2906 ATAGGTTCCAGGCTTGCTGTAATATACCAGATATAGCAAACTCTGGGACAGGAATAAT 2965
|||:|||||
QY 767 IleSerPheIleIleThrGlyTrpGluMetThrPheIleLeuSerIleAlaProValLeu 786
|||:|||||
Db 2966 ATATCCTTCATCTATGTTGGCAACTAAACACTGTTACTCTTAGCAATTTGACCCATCAT 3025
|||:|||||
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
|||:|||||
Db 3026 GCAATAGCAGGAGTGTGTGAAATGTAATGTTGTCGAGCAACGACTGGAAGATAGAAA 3085
|||:|||||
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleAtgThrIleVal 826
|||:|||||
Db 3086 GAACAGAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGGAACCGTTGT 3145
|||:|||||
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetThrGluGluMetLeuGlnThrGlnHis 846
|||:|||||
Db 3146 TCCTTGACTCAGGACGAGAAAGTTTGAACATATGATGTCAGAGTTTGCAGGTACCATAC 3205
|||:|||||
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
|||:|||||
Db 3206 AGAACTCTTTGAGAAAGCACACATCTTGGAAATTACATTTTCCTTACCCAGGCATG 3265
|||:|||||
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
|||:|||||
Db 3266 ATGTATTTTCTATGCTGGATGTTCCGGTTTGGAGCTTACTTGGTGCGCACATAAACTC 3325
|||:|||||
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
|||:|||||
Db 3326 ATGAGCTTTCAGGATGTTCTGTAGTATTTTACGCTGTTGCTTGTGTGCGCCGCTG 3385
|||:|||||
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
|||:|||||
Db 3386 GGGCAAGTCAGTTCATTTGCTTCCTGATGCTATGCCAAAGCCAAATATCAGAGCCCATC 3445
|||:|||||
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
|||:|||||
Db 3446 ATCATGATCATTTGAAAACCCCTTTGATTGACAGCTACAGCAGGAAGCGCTAATCCG 3505
|||:|||||
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgPro 966
|||:|||||
Db 3506 AACACATTGGAAGGAATGTCCATTTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3565
|||:|||||
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
|||:|||||
Db 3566 GACATCCAGTGCTTCAGGACTGAGCTGGAGTGGAAGGCCAGAGCTGGCTCTG 3625
|||:|||||
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
|||:|||||
Db 3626 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTGTCCAGCTCTCTGGAGCGGTTCTACG 3685
|||:|||||
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
|||:|||||
Db 3686 CCCTTGGCAGGGAAGTGCTCTGTATGGCAAGAAATAAAGCCAGCAATGTTCACTGG 3745
|||:|||||
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
|||:|||||
Db 3746 CTCGAGCACACCTGGGCATCGTGTCCAGGAGGCCATCTCTGTTGACTGCAGCATTTGCT 3805
|||:|||||
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
|||:|||||
Db 3806 GAGAACATTGCTTATGGAGACAAACAGCCGGGTGTGTGCACAGGAAGAGATTGTGAGGCA 3865
|||:|||||
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
|||:|||||
Db 3866 GCAAGGAGGGCCACATACATGCTTTCATGAGTCACTGCGCTTAATAATAATAGCACTAAA 3925
|||:|||||
```

```
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
|||:|||||
Db 3926 GTAGGAGACAAAGGAACCTCAGCTCTCTGGTGCCAGAAACGCAATTCCTAGCTCGT 3985
|||:|||||
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
|||:|||||
Db 3986 GCCCTTGTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCACGCTCAGCTCTGGATAC 4045
|||:|||||
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
|||:|||||
Db 4046 GAAAGTGAANAAGTGTCTCCAGAAGCCCTGGCAAAAGCCAGAGAGCCGACCTGCATT 4105
|||:|||||
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
|||:|||||
Db 4106 GTGATTGCTCACCGCTGTCCACCATCCAGAAATGCAGACTTAATAGTGTGTTTCAGAA 4165
|||:|||||
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
|||:|||||
Db 4166 GGCAGAGTCAAGGAGCATGGCAGCATCAGCTGCTGGCACAGAAAGGCATCTATTTT 4225
|||:|||||
QY 1187 LysLeuValAsnAlaGlnSer 1193
|||:|||||
Db 4226 TCATGCTCAGTGTCCAGGCT 4246
|||:|||||
RESULT 13
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-02CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4189
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-5
Alignment Scores:
Pred. No.: 2,71e-308 Length: 4189
Score: 3526.50 Matches: 678
Percent Similarity: 74.74% Conservative: 242
Best Local Similarity: 55.08% Mismatches: 262
Query Match: 58.66% Indels: 49
DB: 10 Gaps: 10
US-09-873-409-6 (1-1195) x US-09-866-866A-5 (1-4189)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
|||:|||||
Db 148 ATGATTCTGGGAACCTCTCGCTGCTATTATCCATGGAACATTACTTCCCTCTTGTATGCTG 207
|||:|||||
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
|||:|||||
Db 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTATTCTGCCAAGCATT 267
|||:|||||
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
|||:|||||
Db 268 ACTAATCAAGTGGACCCCAACAGTACTCTCATCATCAGCAACAGCAGCTCTGGAGGAAGAG 327
|||:|||||
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAla 56
|||:|||||
```

Db 328 ATGCCCATATACCCCTAC-----TATTACACCGGGATTGGTGTGTGTG 372  
 Qy 57 LeuilePheGlyTyrileGlnileSerLeuTriPheIleThrAlaalaArgGlnThrLys 76  
 Db 373 CTCATATGTCCTACATCCAGGTTTCACITTTGGTGGCTGGCAGCTGGAACACAGATACAC 432  
 Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
 Db 433 AAGATTAGGCAGAAAGTTTTCATGCTATATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492  
 Qy 97 CysAspIleGlyLeuLeuThrArgMetThr--AspIleAspLysIleSerAspGly 115  
 Db 493 CATGATTGTTGGGAGCTCAACACCGGCTCACAGATGATGTCTCCAAATAATTAATGACGA 552  
 Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
 Db 553 ATTGTTGACAAANTGGGATGTTTTTCAGTCCATAACACACATTTTAGCCGGTTTATC 612  
 Qy 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
 Db 613 ATAGGATTTATAGTGGTTGGAAGCTAAACCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672  
 Qy 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
 Db 673 GGATTCATCTCTTTGGGCAAAAGGTATTGACTTCAATTTACTAATAAGGAACCTCCAG 732  
 Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195  
 Db 733 GCTTATGCAAAAGCTGAGCGATGTTCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792  
 Qy 196 AlaPheArgAlaGlnGlyLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215  
 Db 793 GCCTTTGGAGGCAACACAGAGGAACTTGAAAGGTACAATAAATAATTAGAAAGACTAAA 852  
 Qy 216 AppPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235  
 Db 853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCCGATAGGCATTCCTACCTGTTG 912  
 Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu 255  
 Db 913 GTCTATGCATCATATGCATGGCATTCGTGGTATGGACATCCTTGGTCTCTCAATGAA 972  
 Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyr 275  
 Db 973 -----TATTCTATTGGAGAAGTGTACTGTCTTCTCTCTATTGTTGGGGACTTTT 1026  
 Qy 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295  
 Db 1027 AGTATTGGACACTTGGCCCCCAACATAGAGCCTTTGCAAAACGCAGCGGGCAGCCTTT 1086  
 Qy 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315  
 Db 1087 GAATCTTCAAGTAAATGTAACGAGCCNAGCATTTGACAGCTTCTCAACAAAGGGCTAC 1146  
 Qy 316 LysProGlySerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335  
 Db 1147 AAACACAGACAGTATAATGGGAACCTTAGAGTTTAAAAATGTTCCACTTCAACTACCCATCG 1206  
 Qy 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355  
 Db 1207 AGAAGCGAAATTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGCGGCAGACGGGTG 1266  
 Qy 356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375  
 Db 1267 GCCTTGGTTGGCAACAGTGCTGTGAAAAAGCACAACTGTCCAGCTGTATGACAGGGCTC 1326  
 Qy 376 TyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395  
 Db 1327 TAGACCCCTCGAGGGCGTGGTCAGTATCGACGGACAGACATCAGACCATCAATGTG 1386  
 Qy 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThr 415  
 Db 1387 AGGTATCTGAGGGAGATCATTTGGTGTGTGAGTCAGAAACCTGTGTGTTTGGCCACGACG 1446

Qy 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435  
 Db 1447 ATGCCCGAGAACATTTCGTATGGCGAGAAGATGTCCACCATGATGAGATTGAGAAAGCT 1506  
 Qy 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455  
 Db 1507 GTCAAGGAACCAATGCCCTATGACTTCATCAATAAACTGCCCCACCAATTTGACACCTG 1566  
 Qy 456 ValGlyGluLysGlyValaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475  
 Db 1567 GTTGTGAGAGAGGGCGCGAGCTGAGTGGGGACAGAAACAGAGAATCCCATTTGCCCGG 1626  
 Qy 476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495  
 Db 1627 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTGGAGCGCCACCTCAGCCCTGGATACA 1686  
 Qy 496 GluSerLysSerAlaValaGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIle 515  
 Db 1687 GAAAGTGAAGCTGTGTCAGCGCCGACCTGGATAAGGCTAGAGAAGCGCGACCACTT 1746  
 Qy 516 ValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAsp 535  
 Db 1747 GTGATAGCTCATCGCTTGTCTACAGTTCTGTAATGTGACGTCACTGCTGTTTGTATGTG 1806  
 Qy 536 GlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 555  
 Db 1807 GGTGTCATTTGCGACCAAGGAATCATGATGAGCTCATGAGAGAAAGGCGCATTTACTTC 1866  
 Qy 556 SerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 575  
 Db 1867 AAACTGTGTCAGACAG---ACTAGAGGAATGAAATTTGAACACGAGAAATAATGCTTAT 1923  
 Qy 576 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAsp 595  
 Db 1924 GGATCCAGAGTGACACTGATGCTTCTGAACCTGACTTCAGAAGAAATCC---AAATCACCT 1980  
 Qy 596 PheIleAspLysAla-----GluGluSerThrGlnSerLysGluIle 609  
 Db 1981 TTAATAAGGAGATCAATTTTACAGAAAGTGTCCACAGAAACCAAGACCAAGAGAGACTT 2040  
 Qy 610 SerLeu-----ProGluValSerLeuLysIleLeuLys 621  
 Db 2041 AGTATGAAAGAGCGTGTGATGAAGATGTGCCTCTGTTTCTCTTTGGCGGATCTTAAAT 2100  
 Qy 622 LeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 641  
 Db 2101 CTAAATCTAAGTGAATGGCTTATTACTTGTGTGGCTACTTTGCGCTGTATTATAATGGG 2160  
 Qy 642 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---Asn 660  
 Db 2161 TGCATACACAGAGTGTTCCTCATAGTATTTTCAAGGATGTAGGGGTTTTTCAAGAGAT 2220  
 Qy 661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680  
 Db 2221 GATGACCATGAAACTAAACGACAGAAATTTGTAATTTGTTTCTCTGTTCTTCTGTTATG 2280  
 Qy 681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700  
 Db 2281 GGGTGATTTCTTTTGTATATATTCTTTTCAGGGGCTTCACATTTGGCAAGCCGAGAG 2340  
 Qy 701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720  
 Db 2341 ATCTCCACCAAGGAGTCCGATACATGTTTTCNAATCCATGCTGACAGAGATATAGC 2400  
 Qy 721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740  
 Db 2401 TGGTTCGATGACATAAAGAACAGCACTGGCTCACTGACCAACAGGCTCGCCAGTGATGCT 2460  
 Qy 741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760  
 Db 2461 TCTAGTGTAAAAGGGGCGATGGCGCCAGGCTGTGTTAGTTATCCAGAAATGTAGCAAAC 2520

```
QY 761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 780
Db 2521 CTGGGACAGGAGTCATCTCTCTCTGATCTATGGTGGCGGACGCTTCTACTTGTGA 2580
QY 781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800
Db 2581 GTAATTATACCGCTCAITGATTGGCGGAATTATTGAATGAAGCTGTTGTCTGGCCAA 2640
QY 801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 820
Db 2641 GCCTTGAGGACACAAACAGCTTGAGTCTCTGGGAAGATTGCTACAGAAGCAATTGAA 2700
QY 821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840
Db 2701 AACTTCGGCACTATTGTCTCTTTGACTCGGGAGCAAGATTTGAAACCATGTATGCCAC 2760
QY 841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860
Db 2761 AGCTTGCGAGTACCATACAGAAATGCGATGAAGAAAGCACACGTGTTTGGGATCAGCTTC 2820
QY 861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880
Db 2821 TCCTTACCAGCCGATGATATTTCTATGCTGCTGTTTCCGTTCCGTTGCTGCTAC 2880
QY 881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900
Db 2881 TTGCTGGCACAAACACTCATGACTTTTGAATAATGTTATGTTGTTATTTCTGCTGTTC 2940
QY 901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920
Db 2941 TTTGGTCCCATGGCAGCTGGGAATACTAGTTCAATTTGCTCCTGACTATGCGAAGCCAAA 3000
QY 921 SerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSer 940
Db 3001 GTATCAGCATCTCATATCATCAGATCATTTGAGAAACCCCTGAGATTGACAGCTACAGC 3060
QY 941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960
Db 3061 ACAGAGGCTTGAAGCTACTCTCTTGAAGAAATGTAATAATTTAATGAGTCCAGTTT 3120
QY 961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980
Db 3121 AACTATATCCACCCGACCAATCCCATGCTTCCAGGGCTGAGCCCTCGAGGTGAAGAAG 3180
QY 981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000
Db 3181 GCCCAGACGTTGGCCCTGCTGGGACAGCTGGCTGTGGGAGAGCACAGCTGCTCCAGCTG 3240
QY 1001 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 1020
Db 3241 CTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTCTTAGATGCCAAAGAAATAAG 3300
QY 1021 GluLeuAsnValGlnThrLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040
Db 3301 CAACTGAATGTCCAGTGGCTCCGAGCTCACCTTGGCAATGTGTCCAGGAGCCCATCTC 3360
QY 1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060
Db 3361 TTTGACTGCGACATGTGAGAGAAATCGCTTATGAGACACACCGCGCGCTGTCTCAT 3420
QY 1061 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 1080
Db 3421 GAGGAGATTGTGAGGCGAGCCAGGAGGCCACATCCACCATGTTTCATCGCTCCT 3480
QY 1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGln 1100
Db 3481 GATAAATACAAACACAGATGAGGAGACAAAGGCACTCAGCTGTGCGGTGGCGAGACAG 3540
QY 1101 ArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120
Db 3541 CGCATCCCATCGCAGCTGCTCGTACAGACGCTCACAATTTTACTTCTGAGCAGAGCA 3600
QY 1121 ThrSerAlaLeuAspAsnArgSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140
```

```
Db 3601 ACATCAGCTCTGGATACAGAAAGTGAAGAGTGTCCAGAAACGCTGGACAAAGCCAG 3660
QY 1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160
Db 3661 GAAGGCCGACCTGCTGATTTGATGCTCATCCGCTGTCCACCATCCAGAACGGGACTTG 3720
QY 1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180
Db 3721 ATCGTGTGATTGAGAACGGCAAGTCAAGGACGACCGCACCCACGACGCTGCTGGCG 3780
QY 1181 AsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db 3781 CAGAAGGCGATCTACTTCTCAATGTGTCCAGGCT 3813
RESULT 14
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US2002010244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIF2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 4788
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-7
Alignment Scores:
Pred. No.: 7,58e-308 Length: 4788
Score: 3522.50 Matches: 668
Percent Similarity: 74.74% Conservative: 252
Best Local Similarity: 54.26% Mismatches: 264
Query Match: 58.59% Indels: 47
DB: 10 Gaps: 8
US-09-873-409-6 (1-1195) x US-09-866-866A-7 (1-4788)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGCTGGTGGAACTCTGGCTGCTATATCCATGGAGTGGCGCTCCCACTTATGATGCTG 207
QY 21 ValLeuGlyCyluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
Db 208 ATCTTTGGTGACATGACAGATAGCTTTGGCAAGTGTAGGAAACGCTCTCTAAAAACAGTACT 267
QY 39 -----Thr 39
Db 268 AATATGAGTGAGCGCGATGAGAGAGCATGTTTGGCAAACTGGGAGGAGAAATGACCAAG 327
QY 40 TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 328 TAGCCTAC-----TATTACCGGAATTTGGTGTGCTGCTCATAGTT 372
QY 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 373 GCCTACATCCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAGATACACAAGATCAGG 432
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 433 CAGAAATTTTTCATCTATAATGAATCAGGATAGGCTGGTGTGATGTGCATGACGTT 492
```



QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
DB 493 GGGAGCTCAACACCGGCTCACAGATGATGTTTCCAAATAATGAAGCAATTGGTGAC 552  
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
DB 553 AAAATCGGAATGTTCTTCAGAGCAATGGCAACATTTTTGGTGGTTTATAATAGGATTT 612  
QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
DB 613 ACCGTGGTGGAGCTAACCTTGTGATTTGGCCATCAGCCCTGTTCTGGACTGTCA 672  
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
DB 673 GCTGGTATTTGGCAAGATATTGCTCTCATTTACTGATAAGCACTCCATGCTTATGCA 732  
QY 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
DB 733 AAAGCTGGAGCAATGCTGGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTCGTTTGA 792  
QY 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
DB 793 GCACAAAGAGGAACCTTGAAGGTACAATAACAATTTGAAGAGCTAAAGGCTGGGG 852  
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238  
DB 853 ATAAAGAAAGCTATCACGCCCAACATCTCCATGGGTGACGTTTCTCTCTATGCA 912  
QY 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258  
DB 913 TCATATGCTCTGGCATCTCGTATGGGACTTCCTTGGTCACTCCAAAGAA-----TAC 966  
QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGly 278  
DB 967 TCTATTGGCAAGTGTCTACTGTCCTTCTTTCGGTGTAAATGGAGCATTCAGTGTGGA 1026  
QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298  
DB 1027 CAGGCATCTCCAAATATTGAAGCTTCGCCAATGCACGAGGACAGCTTATGAAGTCTTC 1086  
QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
DB 1087 AAAATAATTGATAAATAAGCCAGTATAGACAGCTTCTCAAGAGATGGGCACAAACAGAC 1146  
QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
DB 1147 AACATACAAGGAATCTGGAATTTAAGAATATTCACTTACCTTACCCATCTCGAAAGAA 1206  
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358  
DB 1207 GTTCAGATCTTGAAGGCTCAATCTGAAGTGAAGCGGACAGACGCTGGCCCTGGTT 1266  
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378  
DB 1267 GGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAAAAGGCTCTACGACCCC 1326  
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398  
DB 1327 CTAGATGGCATGTGTGATCTCGACGGACAGGACATCAGAACCATCAATGTGAGGTATCTG 1386  
QY 399 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 418  
DB 1387 AGGAGATCATTTGGTGGTGTGACTCAGGAACCTGTGCTGTTCCACCACGATCGCCGAG 1446  
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438  
DB 1447 AACATTCGCTATGGCCGAGGAAGATGTCAACCATGGATGAGATTGAGAAGCTGTCAAGGAA 1506  
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458  
DB 1507 GCCAATGCCTATGACTTCATGAAACTGCCCCACCAATTTGACACCCCTGGTGGTAG 1566

QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478  
DB 1567 AGAGGGGGCCAGCTAGTGGGGGAGAGAAACAGAGAAATGCCATTTGCCGGGGCCCTGGTC 1626  
QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498  
DB 1627 CCATATCCCAAGATCTTGTGTGGACGAGGCCACTCAGCCCTGGATACAGAAAGTGAA 1686  
QY 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518  
DB 1687 GCTGTGTTTCAGCCGCACTGGATGAAGCTAGAGAAGCGCGACCACTTGTGTAGTACT 1746  
QY 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538  
DB 1747 CATCGTGTCTACCGCTCGTAATGTCGACGTCATTCGCTGGTGGTTTGAATGGTGTGTCATT 1806  
QY 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558  
DB 1807 GTGGAGCAAGGAATCATGATGCTCATGAGAAAGAGGCATTTACTTCAAACTTGTC 1866  
QY 559 MetSerGln-----AspIleLysLysAlaAspGlu 568  
DB 1867 ATGACACAGACAGCAGGAATGAAATTTGAATTTAGGAATGAAGCTTGTAAATCTAAGGAT 1926  
QY 569 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 588  
DB 1927 GAAATGTATAATTTAGACATGCTCTCAAAAGATTCCAGGATCCAGTCTAATAAGAAGA--- 1983  
QY 589 ValLysSerIleLysSerAspPheIle-----AspLysAlaGluGluSer 603  
DB 1984 ---AGATCAACTTCGCAAAAGCATCTGTGGACCACTGATGACCAAGAGGAGCTTAGTACC 2040  
QY 604 ThrGlnSerLysGlyLeuSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 623  
DB 2041 AAAGAGGCCCTGGATCAAGATCTCACTCCAGCTCTCTTTGGCGGATCTCTGAAGTTGAAT 2100  
QY 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643  
DB 2101 TCAACTGAATGGCTTATTTTGTGGTGTATATTCTGTGCCATAATAAATGAGGCTTA 2160  
QY 644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 662  
DB 2161 CAGCCAGCATCTCCGTAATAATTTCAAAAGTTGTAGGGGTTTTTACAATGGTGGCCCC 2220  
QY 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682  
DB 2221 CTGAAACCCAGCGGCGCAGACAGCACTTGTTCCTCTGTTCTGATCTCTGGGATC 2280  
QY 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702  
DB 2281 ATTCTTTTCAATACATTTTCTTCTAGGCTTTCACATTTGGCAAAAGCTGAGAGATCTC 2340  
QY 703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe 722  
DB 2341 ACCAGCGATCCGATACATGTTTCAATCCATGCTCAGACAGGATGTGAGCTGGTTT 2400  
QY 723 AspGlyLysGluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742  
DB 2401 GATGACCTTAAACACACCGGACGACTGACCAAGGCTCGCCAAACGATGCTGCTCAA 2460  
QY 743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762  
DB 2461 GTMAAGGGGCTACAGGGTCTAGGCTTGTGTTGATTTTCCAGAAACATAGCAATCTTGG 2520  
QY 763 LeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIle 782  
DB 2521 ACAGGAATCATCATATCCCTAATCTATGCTGGCAACTTAACACTTTTACTCTTAGCAAT 2580  
QY 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802  
DB 2581 GTACCCATCTTCGATACCTGGAGTGGTTGAATGAAATGTTGTCTGGCAGGACACTG 2640  
QY 803 LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822



```
Db 2641 AAAGATAAGAAAGCAACTAGAAAGTTCTCGAAAGATTGCTACGGAAGCAATTAAGAACTTC 2700
Qy 823 ArgThrIleValSerLeuThrArgGluLeuAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2701 CGCACTGTTCTCTTTGACTCGGGAGCAGAAAGTTGAAACCACTGATGCGCCAGAGCTTG 2760
Qy 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db 2761 CAGATACCATACAGAAATCGATGAGAAAGACACAGTGTTCGGATCAGCTTCTCCTTC 2820
Qy 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2821 ACCAGGCCCATGATGATTTTCTTCTATGCTGCTGTTGTTCCGGTTCGGTCTACTTGGTG 2880
Qy 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2881 ACAACAACACTCAGCTTTTGAAGATGTTCTGTAGTATTCTCAGCTATGTTCTTGGT 2940
Qy 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 2941 GCATGCCATGGGGAGGTCACTTCTCGTCTCGTACATTCGAAAGCAACAGTGTCA 3000
Qy 923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db 3001 GCATCCACATCATCAGGATCATTTGAGAAACCCCGAGATTGACAGCTACAGCAGCAA 3060
Qy 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 962
Db 3061 GGCTTAAGCCGATATGTTGGAGGAAATGTGCATTTAGTCGAGTCGTGTTCAACTAT 3120
Qy 963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db 3121 CCACCCGACCCAGCATCCAGTGTTCAGGGGCTGAGCCCTTGAGGTGACAGAGGCCAG 3180
Qy 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db 3181 ACGTGGCCCTGGTGGGACAGTGGCTGGGGAAGACAGTGGTCCAGCTGCTCGAG 3240
Qy 1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3241 CGTTCTACGACCCCATCGGTGGATCATGTTCTAGATGCCAAGAAATTAAGCAACTG 3300
Qy 1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db 3301 AATGTCATGGTCTCCAGCACAGCTGGGATGTGTGCCAAGAGCCCATCTCTTTGAC 3360
Qy 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3361 TGCAGCATCGCAGAGAACATTGGCTACGGAGACAACAGCCGGTCTGTCTTTATGAGGAG 3420
Qy 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3421 ATTGTGAGGGCGCAGCAGAGGCGCAACATCCACCGATTCTATCGCACTCGCTACCTGATAA 3480
Qy 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3481 TACACACCAAGTAGAGACAAGGCACTCAGCTGTCGGGTGGGAGAGAGCGGATC 3540
Qy 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 1122
Db 3541 GCCATCGCAGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGCGAGCAACATCA 3600
Qy 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3601 GCTCTGATACAGAAAGTGAAGGTTGTCCAGAGCGCTGGCAACAAGCCAGCGGAGGC 3660
Qy 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3661 CGCACCTGCAATGTGATCGCTCCAGGCTGTCACCATCCAGAACCGGCACTTGATCGTG 3720
Qy 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
```

```
Db 3721 GTGATTTCAGAACGGCAAGGCTCAAGGAGCAGCGCACCCACAGCAGCTGCTGGCGCAGAAG 3780
Qy 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3781 GGCATCTACTTCTCAATGGTCACTGTGCAAGGCT 3813
RESULT 15
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299
```

```
Alignment Scores:
Pred. No.: 2,38e-306 Length: 3924
Score: 3504.50 Matches: 676
Percent Similarity: 74.43% Conservative: 235
Best Local Similarity: 55.23% Mismatches: 276
Query Match: 58.29% Indels: 37
DB: 10 Gaps: 8
```

```
US-09-873-409-6 (1-1195) x US-09-880-107-2299 (1-3924)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20
Db 201 ATGTGCTGGTACCATCATGCGCATAGCTCAGGATCAGGTCTCCCTCATGATGATA 260
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 261 GTATTTGGAGAGATGACTGCAAAATTTGTTGAT-----ACTGCGGAAACTTC 308
Qy 41 SerPhe-----PheArgLeuThrLeu----- 47
Db 309 TCCTTTCCAGTGAACATTTTCCTTGTGCTGCTAAATCCAGGCAAAATTCGGAAGAGAA 368
Qy 48 -----TyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 369 ATGACTAGATATGCATATTACTACTCAGGATTTGGTGTGGAGTCTTGTGCTGCTAT 428
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 429 ATACAAGTTTCATTTTGGACTTTTGGCAGCTGGTTCGACAGATCAGGAAATTTAGCGAGAAG 488
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 489 TTTTTCATGCTATTCTACGACAGGAAATAGGATGTTTGACATCAATGACACCACTGAA 548
Qy 102 LeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 549 CTCATATCCGGGTAAACATGACATCTCCAAATCAGTGAAGAAATTTGGTGCACAGGTT 608
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
```

Db 609 GGAATGTTCTTTCAAGCAGTAGCCACCGTTTTTTCAGGAGTTCATAGTGGGATTCATCAGA 668  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 669 GGATGGAAGCTCACCTTTGTGATAATGGCCATCAGCCCTATTCTAGGACTCTCTGCGAGCC 728  
Qy 161 AlaCysSerArgMetValIleSerLeuSerIleArgThrValIleAlaPheArgAlaGln 180  
Db 729 GTTTGGCCAAAGATCTCTCGCATTTAGTGCACAAAGAACTAGCTCTTATGTCAAAAGCA 788  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 789 GCGCGCGTGCACAAAGAGGCTCTGGGGCCATCAGACTGTGATAGCTTTTCGGGGGCCAG 848  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 849 AACAAAGAGCTGGAAGGATATCAGAAACATTTAGAAAAATGCCAAAGAGATTGGAATTTAA 908  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 909 AAAGCTATTTCAGCAACATTTCCATGGGTATTGCCCTTCCTGTTAATATATATGATCATAT 968  
Qy 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 969 GCACCTGGCCTTCGGTATGATCCACTCTAGTCTATCAAAAGAA-----TATACTATT 1022  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 1023 GGAATGCAATGACAGTTTTTTTCAATCCTAATTTGGAGCTTTTCAGTGTGGCCAGGCT 1082  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1083 GCCCCATGATTGATGCTTTTGGCAATGCAAGAGGAGCAGCATATGTGATCTTTGATATT 1142  
Qy 301 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1143 ATTGATTAATAATCCTAAATTTGACAGTTTTTTCAGAGAGGAGCACAACCCAGACGATC 1202  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1203 AAAGGGAATTTGGAGTTCAATGATGTTTCACTTTTTCACCTTCTCGAGCTAACGTCAG 1262  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1263 ATCTTGAAGGGCTCAACCTGAAGGTGCAGAGTGGCGACAGCGTGGCCCTGGTGGGAAGT 1322  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1323 AGTGGCTGTGGGAAGAGCACACCGTCCAGCTGATACAGAGGCTCTATGACCCCTGATGAG 1382  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1383 GGCACATTTAACATTGATGGCGAGGATATTAGGAATTTAATGTAACTATCTGAGGGAA 1442  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
Db 1443 ATCATTTGGTGTGTGAGTCAGGAGCGGTGCTGTTTCCACCACCAATTTGCTGAAATATT 1502  
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1503 TGTATTGGCCGTGGAAATGTAACTGATGGATAGATAAAGAGCTGTCAAAAGAGGCCAAC 1562  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
Db 1563 GCCTATGAGTTTATCATGAAATTTACCACAGAAATTTGACACCCCTGGTTGGAGAGAGGG 1622  
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1623 GCCCAGCTGATGGTGGCAGAGCAGAGGATCGCCATTCGCTGCCCTGGTTCGCAAC 1682  
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1683 CCCAAGATCCTTCTGCTGATGAGGCCACGTCAGCATTTGGACATTTGGACAAAAGTGAAGCTGAG 1742

Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
Db 1743 GTACAGGCACTCTGGATAAGCCAGAGAAGCGCGACCACCATTTGTGATGACACACCGA 1802  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 1803 CTGTCTACCGTCCGAAATGCAGATGTCATCGCTGGGTGTTCAGGATGGAGTAATTTGGAG 1862  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal----- 558  
Db 1863 CAAAGNAGCCACAGCAACTGATGAAGAAGAGGGGTGCTACTTCAAACTTGTCAACATG 1922  
Qy 559 -----MetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 571  
Db 1923 CAGACATCAGGAAGCCAGATCCAGTCAGAGAAGATTTGAACTAAATGATGAAAAGCGTCC 1982  
Qy 572 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 591  
Db 1983 ACTAGAATGGCCCAAAATGGCTGGAAATCTCGCCTATTATTAGGCATTTCTACTCAGAAAA 2042  
Qy 592 IleLys---SerAspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIle 609  
Db 2043 CTTAAAAATTCAAATGTGTGAGAGAGCCTTGATGTGGAACCCGATCGACTTGAGCA 2102  
Qy 610 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPhe 629  
Db 2103 AATGTGCCACCACTGCTCTTCTGGAAGGTCTGMAACTGAATAAACAAGATGGCCCTAC 2162  
Qy 630 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 649  
Db 2163 TTTTGTGTGGGAACAGATATGTGCCATTTGCCAATGGGGGCTTCAGCCGCAATTTTCAGTC 2222  
Qy 650 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAsp 669  
Db 2223 ATATTCTCAGACATCATAGCGAATTTTGGACCAGCGATGATGCAGTGAAGCAGCAGAAG 2282  
Qy 670 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 689  
Db 2283 TGCAACATATTCTCTTGATTTCTTATTCTTGGGAATTTATTCTTTTATTACTTCTTTC 2342  
Qy 690 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 709  
Db 2343 CTTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATG 2402  
Qy 710 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThr 729  
Db 2403 GCTTTTAAAGCAATGCTAAGCAGGACATGAGCTGGTTTGTAGCCATATAAAACAGTACT 2462  
Qy 730 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 749  
Db 2463 GGTGCATTTCTACAAGACTTGCACAGATGCTGCCAAGTCCCAAGGAGCCACAGAACCC 2522  
Qy 750 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 769  
Db 2523 AGGTGGCTTTAATGTCACAGAAATATAGCTAACCTTGGAACTGGTATTATCATATCATTT 2582  
Qy 770 IleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 789  
Db 2583 ATCTACGGTGGCAGTTAACCCCTATTGCTATTAGCAGTGTGTTCCAATATTATTCGTGTCA 2642  
Qy 790 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnLeuLys 809  
Db 2643 GGAATGTTGAATGAAATTTGTGGTGGAAATGTCAAAAGAGATATAAAAGAACTGGAA 2702  
Qy 810 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 829  
Db 2703 GCTGTGGAAGATTTGCAACAGAGCAATAGAAAATATTAGGACAGTGTGTGCTTTGACC 2762  
Qy 830 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 849  
Db 2763 CAGGAAGAAAAATTTGAAATCAATGTATGTTGAAAAATTTGTGAGCCTTTACAGGAATTC 2822

Search completed: March 31, 2003, 16:13:21  
Job time : 686.971 secs

```
Qy 850 SerLysLeuAlaGlnIlelleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 869
Db 2823 GTCCAGAGGCACACATCTATGGAATTACTTTAGTAGTATCTCAAGCAATTATGATATTT 2882
Qy 870 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrPro 889
Db 2883 TCCTATGCCGGTTGTTTCGATTGTGGCATATCTCATTTGTAATGGACATATGCCGCTTC 2942
Qy 890 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 909
Db 2943 AGAGATGTTAATCTGGTGTGTTTCTGCAATTGTAATTTGGTCAGTGGCTCTAGGACATGCC 3002
Qy 910 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 929
Db 3003 AGTTTCATTTGCTCCAGCATATGCTAAAGCTAAGCTAGCTGTCAGCCCCACATTATTATGCTG 3062
Qy 930 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 949
Db 3063 TTTGAAAGACACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTT 3122
Qy 950 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 969
Db 3123 GAAGGAAATATAACATTAAATGAAGTCGTCTCACTATCCACCCGAGCAACGTCGCCA 3182
Qy 970 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 989
Db 3183 GTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAGGCCAGACACTAGCCCTGGTGGGCAGC 3242
Qy 990 SerGlyCysGlyLysSerThrSerValGlnLeuGlnAlaArgLeuTyrAspProValGln 1009
Db 3243 AGTGGCTGTGGGAAGACACGGTGTCCAGCTCTCTGGAGCGTCTACGACCCCTTGGCG 3302
Qy 1010 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyrLeuArgSer 1029
Db 3303 GGGACAGTGTCTTCGATGTGTCAGAGCAAGCAAGAACTCAATGTCCAGTGGCTCAGAGCT 3362
Qy 1030 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1049
Db 3363 CAATCGGAATCGTGTCTCAGGAGCCTATCTTATTGACTGCGCATTCGCGAGATATT 3422
Qy 1050 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1069
Db 3423 GCCTATGGACACAACAGCCGGTGTATCACAGGATGAAATTTGTAGTGCAGCCAAAGCT 3482
Qy 1070 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1089
Db 3483 GCCAACATACATCCTTTTCATCGAGACGTTACCCCAAAATATCAAAACAGAGTGGGAGAT 3542
Qy 1090 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1109
Db 3543 AGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATC 3602
Qy 1110 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 1129
Db 3603 AGACACCTCAAAATCCTCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAA 3662
Qy 1130 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1149
Db 3663 AAGGTTTGTCAAGAAGCCCTCGCAAAAGCCAGAGAGCCGACCTGCATTTGTGATTGCT 3722
Qy 1150 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIle 1169
Db 3723 CACCGCTGTCCACCATTCCAGATGCAGACTTAATAGTGTGTTTCAGATGGGAGATC 3782
Qy 1170 LysGluGlnGlyThrHisGlnGluLeuArgAsnArgAspIleTyrPheLysLeuVal 1189
Db 3783 AAGGAGCATGGCAGCATCAGAGCTGCTGGCACAGAAAGGACATCTATTTTCAATGGTC 3842
Qy 1190 AsnAlaGlnSer 1193
Db 3843 AGTGTCAGGCT 3854
```

GenCore version 5.1.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 16155.1 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-6  
Perfect score: 6012  
Sequence: 1 MILGILASLVNGACPLMPL.....QELLNRDIYFKLVNAQSVQ 1195

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09873409/runat\_27032003.115418.19228/app\_query\_fasta\_1.7544  
-DB=GenEmbl -QFMT=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1.1.54887 @runat\_27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.ma.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rodi.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6012	100.0	3621	6	AX339032 Sequence
2	5988.5	99.6	3702	6	AX339031 Sequence
3	5760.5	95.8	3699	6	AX478104 Sequence
4	5159.5	85.8	3177	6	AX339030 Sequence
5	4561.5	75.9	2856	6	AX339028 Sequence
6	3574	59.4	4296	10	CRUPGPI
7	3573	59.4	4304	10	CRUPGPI165
8	3557	59.2	4279	6	AX105082
9	3556.5	59.2	4195	6	AX108656
10	3555.5	59.1	4927	10	AF257746
11	3554	59.1	3860	6	AX322787
12	3554	59.1	4279	6	AX105078
13	3554	59.1	4279	6	AX105080
14	3554	59.1	4646	6	AX391099
15	3552	59.1	4045	12	AF269224
16	3551	59.1	4378	6	E02326
17	3549.5	59.0	3858	4	OAU78609
18	3549	59.0	4279	6	AX105057
19	3549	59.0	4317	4	AF045016
20	3549	59.0	4317	6	AX105059
21	3548	59.0	4186	6	AX108654
22	3545	59.0	3860	6	AX322789
23	3545	59.0	4646	6	AX336420
24	3545	59.0	4646	6	AX336708
25	3545	59.0	4646	6	I49610
26	3545	59.0	4646	9	HUMMDR1
27	3545	59.0	4669	6	I08557
28	3545	59.0	6505	6	AR028671
29	3545	59.0	8630	6	AX012320
30	3545	59.0	8630	6	AX012321
31	3545	59.0	9318	6	AR028672
32	3543.5	58.9	4264	6	AR051647
33	3543.5	58.9	4264	6	AR051650
34	3541.5	58.9	4264	6	AR051650
35	3537.5	58.8	3840	6	AX481416
36	3537.5	58.8	3988	6	AX024454
37	3537.5	58.8	4192	9	AF016535
38	3529	58.7	4280	10	CRUPGPI1
39	3527	58.7	4669	6	AR091275
40	3527	58.7	4669	6	AR203322
41	3526.5	58.7	4189	6	AX322791
42	3526.5	58.7	4298	10	MUSMDR
43	3523	58.6	3934	4	CFA419568
44	3522.5	58.6	3888	10	AY082609
45	3522.5	58.6	4788	6	AX322793

ALIGNMENTS

RESULT 1



Db 1714 CAGGATATTTAAAGAGCTGATGAACAGATGGAGTCAATGACATATTTCTACTGAAGAAG 1773  
Qy 581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600  
Db 1774 ACCAACTCACTTCTCTGCTGCTGTAAGAGCATCAAGTCAGACTTCATTGACAGGCT 1833  
Qy 601 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 620  
Db 1834 GAGGAATCCACCCTCAATGAAGAGATAAGTCTTCTGGAAGTCTCTCTATTAATAATTTTA 1893  
Qy 621 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640  
Db 1894 AAGTTAAACACAGCCTGAATGGCTTTTGGTCTCTGGGACATTTGGCTTCTGTTCCTAAT 1953  
Qy 641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 660  
Db 1954 GGAACCTGTTCCATCAGTATTTCCATCACTTTTGCAGAAATTTATTAACCATGTTTGGAAAT 2013  
Qy 661 AsnAspLysThrThrLeuLysHisPheAlaGluIleTySerMetIlePheValIleLeu 680  
Db 2014 AATGATAAAACACACATTAAGCATGATGACAGAAATTTATTCATGATATTCGTCATTTTG 2073  
Qy 681 GlyValIleCysPheValSerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGlu 700  
Db 2074 GGTGTTATTTGCTTGTCTAGTTATTTTCATGAGGGATTTATTTACGGCAGAGCAGGGAA 2133  
Qy 701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAla 720  
Db 2134 ATTTTAAACATGAGATTAGACACTTGGCTTCAAGCCATGTTATATCAGGATATTGCC 2193  
Qy 721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740  
Db 2194 TGGTTTGATGAAAGGAAACACGACAGCAGGAGCTTGACAAACATATTAGCCATAGATATA 2253  
Qy 741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValIleLeuThrGlnAsnAlaThrAsn 760  
Db 2254 GCACAAATTCAGAGGACACAGGTTCCAGGATGGCGTCTTAACACAAATGCACATAAC 2313  
Qy 761 MetGlyLeuSerValIleIleSerPheIleTyGlyTrpGluMetThrPheLeuIleLeu 780  
Db 2314 ATGGGACTTTCAGTATATCTTCTTTATATATGATGGAGATGACATCTCTGATTCGTG 2373  
Qy 781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800  
Db 2374 AGTATTGCTCCAGTACTTCCGCTGACAGGAATGATTTGAAACCGCAGCAATGACTGATTT 2433  
Qy 801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 820  
Db 2434 GCCAAACAAGATAAGCAAGAACTTAAGCATGCTTGAAGATAGCAACTGAAAGCTTTGGAG 2493  
Qy 821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyGlyGlu 840  
Db 2494 AATATACGTACTATAGTGTATTAACAGGGAAGAAAGCCTTCGAGCAATATGATGAAGAG 2553  
Qy 841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTy 860  
Db 2554 ATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCAGATATTGGAAGCTCTTAT 2613  
Qy 861 AlaPheSerHisAlaPheIleTyPheAlaTyAlaAlaGlyPheArgPheGlyAlaTy 880  
Db 2614 GCATTCAGCCATGCCCTTATATATTTTGCCTATGACAGCAGGGTTTCCGATTTGAGCCTAT 2673  
Qy 881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900  
Db 2674 TTAATTCAGCTGGAGCAATGACCCAGAGGGCATGTTTCATAGTTTTCATGCAATTGCA 2733  
Qy 901 TyGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTySerLysAlaLys 920  
Db 2734 TATGGAGCTATGGCCATCGAAGAAACGCTCTGTTTGGCTCTGAAATATTCGAAGCCAAA 2793  
Qy 921 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 940

Db 2794 TCGGGGGCTGCGCATCTGTTTCCCTTGTGGAAAGAAACCAATATATACACAGCCGCACT 2853  
Qy 941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960  
Db 2854 CAAGAAGGGAAAGCCACACATGTGAAGGGAATTTAGAGTTCGAGAAGTCTCTTTTC 2913  
Qy 961 PheTyProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980  
Db 2914 TTCTATCCATGTCGCCAGATGTTTTCATCTCCGTCGCTTATCCCTCAGTATGAGCGA 2973  
Qy 981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000  
Db 2974 GGAAGACAGTAGCATTTGTGGGAGCAGCGCTGTGGGAAAGACACTTCTGTTCACATT 3033  
Qy 1001 LeuGlnArgLeuTyAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 1020  
Db 3034 CTGCAGAGACTTTATGACCCCGTCAAGGACAAAGTGTCTTTGATGCTGGATGCAAAA 3093  
Qy 1021 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040  
Db 3094 GAATTGAATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAAGAGCTGTGCTC 3153  
Qy 1041 PheAsnCysSerIleAlaGluAsnIleAlaTyArgLysAspAsnSerArgValProLeu 1060  
Db 3154 TTCAACTGCAGCATTTGCTGAGAACATCGCTATGTGTGACACAGCGCTGTGTCATT 3213  
Qy 1061 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 1080  
Db 3214 GATGAGATCAAGAAAGCGCAAAATGCAGCAATATCCATTCTTTTATTTGAAGGCTCTCC 3273  
Qy 1081 GluLysTyAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 1100  
Db 3274 GAGAAATACAAACACACAAAGTTCGAGCTGAAAGGAGCAGCAGCTTCTGCGCGCCAGAA 3333  
Qy 1101 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120  
Db 3334 AGACTAGCTTTTGCAAGGCTCTTCTCCAAAACCCCAAAATTTTATTTGTTGATGAGGC 3393  
Qy 1121 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140  
Db 3394 ACTTCAGCCCTCGATAATGACAGTGAGAGGTGGTTTTCAGCATGCTTGTATTAAGCCAG 3453  
Qy 1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160  
Db 3454 AGGGAAAGACATGCTAGTGTCTCACTCAGGCTCTCTGCAATTCAGAACCGCAGATTG 3513  
Qy 1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180  
Db 3514 ATAGTGGTTCTGCACATGGAAGATAAAGGAACAAGGAACCTCATCAAGAGCTCTTGAGA 3573  
Qy 1181 AsnArgAspIleTyPheLysLeuValAsnAlaGlnSerValGln 1195  
Db 3574 ATCCGACATATATTTTAAGTTAGTGAATGCACAGTCAGTCAGCAG 3618

## RESULT 2

AX339031  
LOCUS AX339031 3702 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 13 from Patent WO0194400.

ACCESSION AX339031

VERSION AX339031.1 GI:18129123

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS Frank, M.H. and Sayegh, M.H.

TITLE A gene encoding a multidrug resistance human p-glycoprotein

homologue on chromosome 7p15-21 and uses thereof

Patent: WO 0194400-A 13 13-DEC-2001.

JOURNAL THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES Location/Qualifiers

1..3702

/organism="Homo sapiens" /db_xref="taxon:9606"									
BASE COUNT	1098 a	708 c	856 g	1039 t	1 others				
ORIGIN									
Alignment Scores:									
Pred. No.:	0	Length:		3702					
Score:	5988.50	Matches:		1195					
Percent Similarity:	97.79%	Conservative:		0					
Best Local Similarity:	97.79%	Mismatch:		0					
Query Match:	99.61%	Indels:		27					
DB:	6	Gaps:		1					
US-09-873-409-6 (1-1195) x AX339031 (1-3702)									
QY	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20						
Db	34	ATGATCTCGGGTATACTGGCATCACTGGTCAATGGAGCCTGCTCTTCTTAATGCCACTG	93						
QY	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40						
Db	94	GTTTTAGGAGAAATGAGTGAATACCTTATTAGTGGATGCTCTAGTCCAACTAACACATAC	153						
QY	41	SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly	60						
Db	154	TCCTTCTTCAGGTTGACCCGTATTATGTTGGAATAGGTGTTGCTGCTTGAATTTTGGT	213						
QY	61	TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys	80						
Db	214	TACATACAGATTTCTCTGTGGATTATAACTGCAGCACAGACCAAGAGGATTCGAAAA	273						
QY	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly	100						
Db	274	CAGTTTTTTCATTCAGTTTGGCACAGACATCGCTGGTTGATAGCTGTGACATCGGT	333						
QY	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle	120						
Db	334	GAACCTTAACACTCGCATGACAGACATTCACAAATCAGTCAGTGTGTTGGAGATAAGATT	393						
QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140						
Db	394	GCTCTGTGTTTCAAAAATGCTACTCTTTTCGATTGGCCCTGGCAGTTGGTTTGGTGAAG	453						
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160						
Db	454	GGCTGGAAACTCACCTAGTACTCTATCCACGCTCTCTCTTATATAGCTTTCAGGGCA	513						
QY	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180						
Db	514	GCATGTTCTAGGATGGTCATCTCATTCACCCAGTAAGCAATTAAGTGCCTATTCCAAAGCT	573						
QY	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200						
Db	574	GGGGCTGTGGCAGAAGAAGTCTTGTCTATCAATCCGAACAGTCATAGCCTTTAGGGCCACG	633						
QY	201	GluLysGluLeuGln-----	205						
Db	634	GAGAAAGAACTTCAAGGCTCTTCTTTTAAATATAACAGATATGCTGGTTTATTTT	693						
QY	206	-----ArgTyrThrGlnAsnLeuLysAsp	213						
Db	694	CCCAGTGGCTACTAAGTTGTTCTGTTTNTTGTAAAGGTATACACAGAAATCTCAAGAT	753						
QY	214	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	233						
Db	754	GCAAGAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTAC	813						
QY	234	PhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsn	253						
Db	814	TTCTTTATGAATGGAACCTATGGACTTGTCTTTTGGTATGGAACCTCTCTTGATTCCTAAT	873						
QY	254	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	273						



QY 634 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 653  
Db 2014 ACATTGGCTTCTGTTCTAAATGGAAGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073  
QY 654 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 673  
Db 2074 ATTATACCATGTTTGGAAATATATGATAAACCACATTAAGCATATGACAGAAATTTAT 2133  
QY 674 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 693  
Db 2134 TCATGATATTTGCTCATTTGGGTGTTATTTGCTTTGTCAGTTATTTTCATGACGGGATTA 2193  
QY 694 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 713  
Db 2194 TTTTACGGCAGACGCGGAAATTTTAAACGATGAGATTAAAGACACTTGGCCTTCAAGCC 2253  
QY 714 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 733  
Db 2254 ATGTTATATCAGGATATTCCTGGTGTGTTGATGAAAGGAAACACAGCAGAGGCTTGACA 2313  
QY 734 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 753  
Db 2314 ACATATTAGCCATAGATATAGCACAAATTCAGGAGCAACAGGTTCCAGGATTGGCGTC 2373  
QY 754 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyr 773  
Db 2374 TTTAACACAAATGCAACTAACATGGGACTTTCAGTTATCATTTCTTTATATATGATGG 2433  
QY 774 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 793  
Db 2434 GAGATGACATTTCTGATTTCTGATTTAGTTCAGTACTTTCGCGTGACAGGAATGATTGAA 2493  
QY 794 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 813  
Db 2494 ACCGACGCAATGACTGGATTGTCACAAAGATAAGACAGAACTTAAGCATGCTGGAAG 2553  
QY 814 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 833  
Db 2554 ATAGCAACTGAAGCTTTGGAGATATACGTACTATAGTGTCTTACACAGGGAAGGCC 2613  
QY 834 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 853  
Db 2614 TTCGAGCAAAATGATGAAGAGATGCTTCAGACTCAACACAGCAAAATACCTCGAAGAAAGCA 2673  
QY 854 GlnIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 873  
Db 2674 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCGCTTTATATATTTTTCCTATGACGA 2733  
QY 874 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 893  
Db 2734 GGGTTTCGATTGGAGCCTATTAAATCAAGCTGGACGATGACCCAGAGGCGCATGTTTC 2793  
QY 894 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913  
Db 2794 ATAGTTTTTACTGCAATTCATATGAGCTATGCGCATGCGCCATCGAAACCGCTGTTTGGCT 2853  
QY 914 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 933  
Db 2854 CCTGAATATTTCCAAAGCCAAATCGGGGCTGCGCATCTGTTTCCCTTGTGGAAGAAAGAA 2913  
QY 934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953  
Db 2914 CCNAATATAGACAGCCGACGTCAGAAAGGGAAGAAAGCCAGACATGTGAGGGAATTTA 2973  
QY 954 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 973  
Db 2974 GAGTTTCGAGAACTCTCTTCTTCTATTCATGTCGCGCCAGATGTTTTCATCTCCCTGGC 3033  
QY 974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993  
Db 3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTTGTGGGAGCAGCGCTGTGGG 3093

QY 994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1013  
Db 3094 AAAAGCACTTCTGTTCAACTTCTGAGAGACTTTATGACCCCGTGCAGGACAAAGTGTCTG 3153  
QY 1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033  
Db 3154 TTTGATGGTGTGGATGCAAAAGAAATTGAATGATACAGTGGCTCCGTTCCCAAAATAGCAATC 3213  
QY 1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1053  
Db 3214 GTTCTCTCAAGAGCCTGTGCTCTTCACTGCAGCATTTGCTGAGACATCGCTATGTGTGAC 3273  
QY 1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaIleHis 1073  
Db 3274 AACAGCCGTGTGGTGCATTAGATGAGATCAAGAGAGCCGCAAAATGACGAAATATATCCAT 3333  
QY 1074 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1093  
Db 3334 TCTTTTATTGAAGGCTCTCCCTGAGAAATACAAACAAAGTTGGACTGAAGAGGACACAG 3393  
QY 1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1113  
Db 3394 CTTTCTGGCGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAACCCAAA 3453  
QY 1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133  
Db 3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAAGGTTGTTTTCAG 3513  
QY 1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153  
Db 3514 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCTTAGTGGTCACTCAGAGGCTCTCT 3573  
QY 1154 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1173  
Db 3574 GCATTCAGAAAGCAGATTGATAGTGGTCTTGCACAATGGAAGATTAAGGACAAAGGA 3633  
QY 1174 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193  
Db 3634 ACTCATCAAGAGCTCTCGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693  
QY 1194 ValGln 1195  
Db 3694 GTGCAG 3699  
RESULT 3  
LOCUS AX478104 3699 bp DNA linear PAT 12-AUG-2002  
DEFINITION Sequence 34 from Patent WO0240541.  
ACCESSION AX478104  
VERSION AX478104.1 GI:22217064  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y.,  
Walia, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L.,  
Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G.,  
Azimzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A.,  
Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E.,  
Policky, J.L. and Kearney, L.  
TITLE Transporters and ion channels  
JOURNAL Patent: WO 0240541-A 34 23-MAY-2002;  
FEATURES Location/Qualifiers  
source 1. 3699  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="incyte ID No: 7472030CB1"  
BASE COUNT 1116 a 707 c 860 g 1016 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 3699  
 Score: 5760.50 Matches: 1156  
 Percent Similarity: 95.87% Conservative: 4  
 Best Local Similarity: 95.54% Mismatches: 9  
 Query Match: 95.82% Indels: 41  
 DB: 6 Gaps: 3

US-09-873-409-6 (1-1195) x AX478104 (1-3699)

QY 1 MetIleuGlylleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 145 ATGATCTGGGTACTACTGACATCACTTTCAATGGAGCCTGCCTTCCTTTAAATGCCACTG 204  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
 DB 205 TGTATAGGAGAAATGAGTGAATAACCTTATTAGTGATGTCTAGTCCACACTAACACACA 264  
 QY 41 SerPhePheArg-----LeuThr 46  
 DB 265 AATTATCAGAACTACTCAGTCTCAAGAGAGCTGAATGAAGATATGACTCTCTTGACC 324  
 QY 47 LeuTyrTyrValGlylleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
 DB 325 CTGTATTATGTTGGAATAGGTGTGCTGCCTTGATTTTGGTTACATACAGATTTCCCTG 384  
 QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgIleGlyGlnPheHisSerVal 86  
 DB 385 TGGATTATAACTGAGCACCACAGACCAAGAGGATTCGAAACAGTTTTTTTCATTAGATT 444  
 QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
 DB 445 TTGGCAGAGACATCGGCTGTTTGTAGTGTGCACATCGTGNACTTAACACTGCGATG 504  
 QY 107 Thr---AspIleAspIlylleSerAspGlyIleGlyAspIlylleAlaLeuLeuPheGln 125  
 DB 505 ACAGATGACATTTGCAAAATCAGTGTGTTATGGAGATAAGATTGCTCTGTGTTTCAA 564  
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIlyleGlyTyrLysLeuThr 145  
 DB 565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGTTTGTGAGGGCTGGAACCTCACC 624  
 QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165  
 DB 625 CTAGTGACTCTATCCAGCTCTCTCTTATATGGCTTCAGCGGCGAGCATGTTCTAGATG 684  
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
 DB 685 GTCATCTCTATTGACCAAGGAATTAAGTGCCTTATTCCAAAGCTGGGGCTGTGGCAGAA 744  
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205  
 DB 745 GAAGTCTTGTTCATCAATCCGAACAGTCATAGCCTTTTAGGGCCCGAGGAAGAAGCTTCAA 804  
 QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225  
 DB 805 AGGTATACAGAAATCTCAAGAATGCAAGATTTTGGCATTAAGAAGCACTATAGCTTCA 864  
 QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245  
 DB 865 AAAGTGTCTCTTGTGTGTGTACTTCTTTTGAATGGAACCTATGGACTTGTCTTTTGG 924  
 QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265  
 DB 925 TATGGAAACCTCTTGATTTCTTAATGGAGAACCTGGATATACATCGGGACTGTTCTGCT 984  
 QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
 DB 985 GTTTCTTTAGTGAATCCATAGTAGTATTATGCAATGGAGCAGAGTCCCTCACTTTGAA 1044  
 QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305  
 DB 1045 ACCTTCGCAATAGCCCGAGGAGTGCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCC 1104

QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325  
 DB 1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTTGAATCCATAGAGAACTGTGGAA 1164  
 QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345  
 DB 1165 TTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAAGGCTCTG 1224  
 QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365  
 DB 1225 AATCTCGGAATTAAAGTCTGGAGAGACAGTCGCTTGGTGGTCTCAATGGCAGTGGGAAG 1284  
 QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385  
 DB 1285 AGTACGGTAGTCCAGCTTCTCGAGAGGTTATATGATCCGGATGATGGCTTATCATGGTG 1344  
 QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405  
 DB 1345 GATGAGAATGACATCAGAGCTTTTAAATGTGCGCATTTATCGACACCATATTGGAGTGGTT 1404  
 QY 406 SerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 425  
 DB 1405 AGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAACAATATCAAGTATGGACGAGAT 1464  
 QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445  
 DB 1465 GATGTGACTGATGAGAGATGGAGAGCAGCAGCAGGAGGAGCAAAATGCGTATGATTTTATC 1524  
 QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465  
 DB 1525 ATGGAGTTCCTTAATAATTTTAAATACATTGTTAGGGGAAAAAGAGAGCTCAATAGTGGGA 1584  
 QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485  
 DB 1585 GGGCAGAAAACAGAGGATCGCAATTGCTGCTTGTAGTTCGAAACCCCAAGATTCTGATT 1644  
 QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505  
 DB 1645 TTAGATGAGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGTTCAAGTGTGACTG 1704  
 QY 506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525  
 DB 1705 GAGAAGCGCAGCAAGAGTCCGACTCAATCTGTGTGTAGCAGACCGACTTCTACTATTGGA 1764  
 QY 526 SerAlaAspIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545  
 DB 1765 AGTGCAAGATTGATTGTGACCCCTAAAGGATGGAATGCTGCGGAGAAAGGAGCGCATGCT 1824  
 QY 546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLys 565  
 DB 1825 GAATTAATGGCAAAACAGGCTTATATTATTCTTGTGTGATGTCACAGGATATTAATAAAA 1884  
 QY 566 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 585  
 DB 1885 GCTGATGAACAGATGGAGTCAATGACATATCTACTGAAAGAAAGCAACCACTCCTTCTCT 1944  
 QY 586 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLeuSerThrGln 605  
 DB 1945 CTGCACCTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAA 2004  
 QY 606 SerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPro 625  
 DB 2005 TCTAAAGAGATAGTCTTCTCTGAAGTCTCTCTATTAAAAATTTTAAAGTTTAAACAAGCCT 2064  
 QY 626 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 645  
 DB 2065 GAATGCCCTTTTGTGCTTCTGGGACATTTGGCTTCTGTTCTTAATGGAACTGTTTATCCA 2124  
 QY 646 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 665  
 DB 2125 GTATTTTCCATCATCTTTGCAAAATTTATAACCATGTTTGGAAATTAATGATAAAACACA 2184

Qy	666	Leu	Ysh	i	s	a	s	p	a	l	a	g	l	l	e	t	y	r	s	e	r	m	e	t	i	l	e	p	h	e	v	a	l	l	e	l	e	u	g	l	y	v	a	l	l	e	c	y	s	p	h	e	685									
Db	2185	T	T	A	A	G	A	G	A	T	A	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2244														
Qy	686	V	a	l	S	e	r	T	y	r	P	h	e	M	e	t	G	l	n	G	l	y	L	e	u	P	h	e	T	y	r	G	l	y	A	r	g	A	l	a	G	l	y	G	l	u	l	e	u	T	h	r	M	e	t	A	r	g	705			
Db	2245	G	T	C	A	G	T	T	A	T	T	T	T	C	A	T	G	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2259									
Qy	706	L	e	u	A	r	g	H	i	s	L	e	u	a	l	a	P	h	e	L	y	s	A	l	a	M	e	t	L	e	u	T	y	r	G	i	n	A	s	p	i	l	e	A	l	a	T	r	p	h	e	A	s	p	G	l	u	L	y	s	725	
Db	2260	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2286								
Qy	726	G	l	u	a	s	n	S	e	r	T	h	r	G	l	y	L	e	u	T	h	r	T	h	r	l	e	L	e	u	a	l	a	l	e	A	s	p	i	l	e	A	l	a	G	l	n	i	e	G	l	n	G	l	y	745						
Db	2287	G	A	A	A	C	A	G	C	A	C	A	G	A	G	C	T	T	G	A	C	A	A	T	A	T	T	A	T	T	A	G	C	A	T	A	G	A	T	A	T	A	G	C	A	T	A	T	A	G	C	A	A	A	T	T	C	A	A	G	A	2346
Qy	746	A	l	a	T	h	r	G	l	y	S	e	r	A	r	g	i	e	G	l	y	V	a	l	L	e	u	T	h	r	G	i	n	A	s	n	A	l	a	T	h	r	A	s	n	M	e	t	G	l	y	L	e	u	S	e	r	V	a	l	765	
Db	2347	G	C	A	A	C	A	G	T	T	C	A	G	A	T	T	G	C	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	2406				
Qy	766	I	l	e	i	S	e	r	P	h	e	i	L	e	t	y	r	G	l	y	T	r	P	h	e	T	h	r	P	h	e	L	e	u	L	e	u	S	e	r	i	L	e	A	l	a	P	r	o	V	a	l	785									
Db	2407	A	T	C	A	T	T	T	T	T	T	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	T	A	2466				
Qy	786	L	e	u	a	V	a	l	T	h	r	G	l	y	M	e	t	I	l	e	G	l	u	T	h	r	A	l	a	M	e	t	T	h	r	G	l	y	P	h	e	A	l	a	S	n	L	y	a	S	p	L	y	s	805							
Db	2467	C	T</																																																											

Db	3187	TGGCTCGTTCGCAAAATAGCAATCGTTCTCTCAAGAGCCTGTGCTCTTCAACTGCAGCATT	3246
Qy	1046	AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu	1065
Db	3247	GCTGAGAACATCGCTTATGTCACAAACGCCGTGTGGTGCCATTAGATGAGATCAAGAA	3306
Qy	1066	AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr	1085
Db	3307	GCCGCAAAATGCAGCAATATCCATTCTTTTATTTGAAGGTCTCTCCCTGAGAAATACAACACA	3366
Qy	1086	GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla	1105
Db	3367	CAAGTTGGACTGAAGAGCACAGCTTCTGCGGCCCAAGAAACAAGACTAGCTATTGCCA	3426
Qy	1106	ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp	1125
Db	3427	AGGCTCTTCTCCAAAACCCCAAAATTTATTGTTGGATGAGGCCACTTCAGCCCTCGAT	3486
Qy	1126	AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys	1145
Db	3487	AATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGTATAAAGCCAGGACCGGAAGACATGC	3546
Qy	1146	LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis	1165
Db	3547	CTAGTGGTCACTCACAGGCTCTCTGCATTTCAAGACGAGATTGTGATAGTGGTTCTGCAC	3606
Qy	1166	AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr	1185
Db	3607	AATGGAAGATTAAGGAACAAGAACTCATCAAGAGCTCTTGAGAAATCGAGACATATAT	3666
Qy	1186	PheIysLeuValAsnAlaGlnSerValGln 1195	
Db	3667	TTTAAAGTTAGTGAATGCACAGTCAGTGCAG 3696	
RESULT 4			
AX339030			
LOCUS	AX339030	3177 bp	DNA linear PAT 09-JAN-
DEFINITION	Sequence 12 from Patent WO0194400.		
ACCESSION	AX339030		
VERSION	AX339030.1	GI:18129122	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomom		
REFERENCE	1		
AUTHORS	Frank, M. H. and Sayedh, M. H.		
TITLE	A gene encoding a multidrug resistance human p-glycoprotein		
JOURNAL	homologue on chromosome 7p15-21 and uses thereof		
FEATURES	Patent: WO 0194400-A 12 13-DEC-2001;		
source	THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)		
	Location/Qualifiers		
	1..3177		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	970 a 601 c 735 g 870 t	1	others
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	3177
Score:	5159.50	Matches:	1031
Percent Similarity:	97.45%	Conservative:	0
Best Local Similarity:	97.45%	Mismatches:	0
Query Match:	85.82%	Indels:	27
DB:	6	Gaps:	1
US-09-873-409-6 (1-1195) x AX339030 (1-3177)			
Qy	165	MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla	184
Db	1	ATGTCATCTCATTTGACCACTAAGAAATTAAGTGCCTATTCCAAAGCTGGGCTGTGCCA	60



Db 2221 GGAGCCTATTAAATCAAGCTGACGAATGACCCCGAGAGGGCAATGTTTCATAGTTTTTACT 2280  
Qy 898 AlaileAlaTyrglyAlaMetAlaileGlyLysThrLeuValLeuAlaProGluTyrsr 917  
Db 2281 GCAATTTGCATATGAGACTATGGCCATCGGAAACCGCTCGTTTGGCTCTCTGAATATTC 2340  
Qy 918 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAlaileAsp 937  
Db 2341 AAGCCAAATCGGGGCTCGCATCTGTTTGGCTTGTGGAAAGAAACCAATATAGAC 2400  
Qy 938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957  
Db 2401 AGCCGAGTCAGACGGAAGAAAGCCAGACACATGTGAGGGAAATTTAGAGTTTCAGAA 2460  
Qy 958 ValSerPhePheTyProCysArgProAspValPheileLeuArgGlyLeuSerLeuSer 977  
Db 2461 GTCTCTTTCTTCTATCATGTCGCCCAGATGTTTTCATCTCGTGGCTTATCCCTCAGT 2520  
Qy 978 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 997  
Db 2521 ATTGAGCGAGGAAGACAGTAGCATTTGTTGGGAGCAGCGGCTGTGGGAAAGCACTTCT 2580  
Qy 998 ValGlnLeuLeuGlnArgLeuTyArgProValGlnGlnGlnValLeuPheAspGlyVal 1017  
Db 2581 GTTCAACTTCTGAGAGACTTTATGACCCCGTCGACAGGCAAGTGTGTTGATGGTG 2640  
Qy 1018 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnileAlaileValProGlnGlu 1037  
Db 2641 GATGCAAGAGATTGATGTACAGTGGCTCCGTTCCCAATAGCAATCGTTCTCAGAG 2700  
Qy 1038 ProValLeuPheAsnCysSerileAlaGluAsnileAlaTyrglyAspAsnSerArgVal 1057  
Db 2701 CCTGTGCTCTTCAACTGCAGCATTTGTGAGAACATCGCCTATGTGACACACGCCGTGT 2760  
Qy 1058 ValProLeuAspGluLeuLysGluAlaAlaAsnAlaAlaAsnileHisSerPheileGlu 1077  
Db 2761 GTGCCAATTAGATGAGATCAAGAGCGCAATGCGCAAAATGCGCAAAATATCCATTTTATGAA 2820  
Qy 1078 GlyLeuProGluLysTyrsrAnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1097  
Db 2821 GGTCTCCTTGAGAAATACACACACAGTTGGACTGAAAGGACACAGCTTCTCGCGGC 2880  
Qy 1098 GlnLysGlnArgLeuAlaileAlaArgAlaLeuLeuGlnLysProLysileLeuLeuLeu 1117  
Db 2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTCTCCAAAACCCCAAAATTTTATTTGTTG 2940  
Qy 1118 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1137  
Db 2941 GATGAGGCCACTTCAGCCCTCGATATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGAT 3000  
Qy 1138 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaileGlnAsn 1157  
Db 3001 AAGCCAGGAGGGAGGACATGCTAGTGGTCATCTCAGAGGCTCTCTGCANTTCAGAAC 3060  
Qy 1158 AlaAspLeuileValValLeuHisAsnGlyLysileLysGluGlnGlyThrHisGlnGlu 1177  
Db 3061 GCAGATTGATAGTGGTTCTGCACATGGAAGAGATGAAGGAAACCAAGAACTCATCAAGAG 3120  
Qy 1178 LeuLeuArgAsnArgAspIleTyPheLysLeuValAsnAlaGlnSerValGln 1195  
Db 3121 CTCTCGAGAAATCGAGACATATATTTTAACTTAGTGAATGCAATGACAGTCAGTGCAG 3174

RESULT 5  
AX339028  
LOCUS AX339028 2856 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 10 from Patent WO0194400.  
ACCESSION AX339028  
VERSION AX339028.1 GI:18129120  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
JOURNAL homologue on chromosome 7p15-21 and uses thereof  
Patent: WO 0194400-A 10 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES location/Qualifiers  
source 1..2856  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 888 a 540 c 652 g 776 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.1e-288 Length: 2856  
Score: 4561.50 Matches: 912  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 75.87% Indels: 18  
DB: 6 Gaps: 1  
US-09-873-409-6 (1-1195) x AX339028 (1-2856)  
Qy 266 ValPhePheSerValileHisSerSerTyrcysileGlyAlaAlaValProHis-PheGI 285  
Db 114 GTTTTCTTTAGTGAATCCATAGCAGTATTTCATTGGAGCAGCAGTCCCTCATTTATT-- 171  
Qy 285 uThrPheAlaileAlaArgGlyAlaAlaPheHisilePheGlnValileAspLysLysPr 305  
Db 172 -----GATNAGAAACC 182  
Qy 305 oSerileAspAsnPheSerThrAlaGlyTyrcysProGluSerileGluGlyThrValGI 325  
Db 183 CAGTATAGATAACTTTTCCACAGCTGATATAAACCTGAATCCATAGAGAACTGTGA 242  
Qy 325 uPheLysAsnValSerPheAsnTyrcProSerArgProSerileLysileLysLysGlyLe 345  
Db 243 ATTATAAAATGTTTCTTCAATATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302  
Qy 345 uAsnLeuArgileLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 365  
Db 303 GAATCTCAGAAATTAAGTCTGGAGACAGTCGCTTGGTCTCANTGGCAGTGGAA 362  
Qy 365 sSerThrValValGlnLeuLeuGlnArgLeuTyrcProAspAspGlyPheileMetVa 385  
Db 363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGATGATGGCTTTATCATGGT 422  
Qy 385 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrcArgAspHisileGlyValVa 405  
Db 423 GGATGAGAAATGACATCAGAGCTTTAAATGTCGGCATTTATCGAGACCATATTGGAGTGGT 482  
Qy 405 lSerGlnGluProValilePheGlyThrThrileSerAsnAsnileLysTyrcGlyArgAs 425  
Db 483 TAGTCAAGAGCTGTTTGTTCGGGACCACTCAGTAAACATATCAAGTATGGACAGA 542  
Qy 425 pAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrcAspPheil 445  
Db 543 TGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGAAACAAATGCGTATCATTTAT 602  
Qy 445 eMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGI 465  
Db 603 CATGGAGTTTCTTAATAAATTTAATACATTGGTAGGGAAAAAGAGAGCTCAATAGTGG 662  
Qy 465 yGlyGlnLysGlnArgGileAlaileAlaArgAlaLeuValArgAsnProLysileLeuill 485  
Db 663 AGGGCAGAAACAGAGATCGCAATGCTCGTCTTAGTTTCGAAACCCCAAGATTCTGAT 722  
Qy 485 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 505  
Db 723 TTTAGATGAGGCTACCTGCTCCCTGATTCAGAAACAGTCACTGTTCAAGCTGCACT 782  
Qy 505 uGluLysAlaSerLysGlyArgThrThrileValAlaHisArgLeuSerThrleAr 525  
|||||

Db 783 GGAGAGGCGAGCAAGGTCGGACTCAACTCGTGGTAGCACCGACTTTCTTACTATTTCG 842  
QY 525 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 545  
Db 843 AAGTGCAGATTTGATGTGACCCCTAAAGGATGGAATGCTGGCGGAGAAAGGACACATGC 902  
QY 545 aGluLeuMetAlaLysArgGlyLeuTyTySerLeuValMetSerGlnAspIleLysLy 565  
Db 903 TGAACATAATGGCAAAACGAGGCTCTATATTATTCATTTGATGATTCACAGGATATAAAA 962  
QY 565 sAlaAspGluGlnMetGluSerMetThrTySerThrGluArgLysThrAsnSerLeuPr 585  
Db 963 AGCTGATGAACAGATGAGTCAATGACATATCTTACTGAAAGAAAGACCACTCACTTCC 1022  
QY 585 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysThrGl 605  
Db 1023 TCTGCACCTCTGTGAGAGCATCAAGTCAGACTTCATTTGACAAAGGCTGAGGAATCCACCCA 1082  
QY 605 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 625  
Db 1083 ATCTAAAGAGATAAGTCTTCTCTGAAGTCTCTCTATTAAAAAATTTTAAAGTTAAACAAGCC 1142  
QY 625 oGluTrpPropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 645  
Db 1143 TGAATGGCCCTTTTGTGGTCTGGGGACATTTGGCTTCTGTTCTAAATGGAACTGTTTCATCC 1202  
QY 645 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 665  
Db 1203 AGTATTTTCCATCATCTTTGCAAAAATATTAACCATGTTTGGAAATATGATTAACCCAC 1262  
QY 665 rLeuLysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPh 685  
Db 1263 ATTAAGCATGATGACAGAAATTTATTCATGATATTCGTCTATTTGGTGTATTGGCTT 1322  
QY 685 eValSerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGluIleLeuThrMetAr 705  
Db 1323 TGTCAAGTATTTTCATGCGAGGATTTATTTACGCGAGACAGCGGGGAAATTTTAAACGATGAG 1382  
QY 705 gLeuArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaTrpPheAspGluLy 725  
Db 1383 ATTAAGACACTTGGCCCTTCAAGCCATGTTATATCAGGATATTGCGCTGGTTGATGAAA 1442  
QY 725 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnI 745  
Db 1443 GGAAAAACAGCACAGGAGGCTTGACAAATATTAGCCATAGATATAGCAAAATTTCAAGG 1502  
QY 745 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 765  
Db 1503 AGCAACAGGTTCCAGGATTTGGCGTCTTAAACACAAAATGCAACTAACATGGGACTTTTCAGT 1562  
QY 765 lIleIleSerPheIleTyGlyTpGluMetThrPheLeuIleLeuSerIleAlaProVa 785  
Db 1563 TATCATTTCTTTATATATGATGGAGATGACATTTCTGATTTCTGATTTGCTCCAGT 1622  
QY 785 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 805  
Db 1623 ACTTGGCGTGACAGGAATGATTTGAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682  
QY 805 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaIleuGluAsnIleArgThrII 825  
Db 1683 GCAAGAACTTTAAGCATCTCGGAAGATAGCAACTGAAAGCTTTGGAGAATATACGTACTAT 1742  
QY 825 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyGluGluMetLeuGlnThrGl 845  
Db 1743 AGTGTCTATTAAACAGGGAAGAAAGCCCTTCGAGCAATGATTAAGAGAGATGCTTCAGACTCA 1802  
QY 845 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyTyAlaPheSerHisAl 865  
Db 1803 ACACAGAAATACCTCGAAGAAAGCACAGATATTATTTGGAAGCTGTTATGCATTTAGCCCATGC 1862  
QY 865 aPheIleTyTyPheAlaTyAlaAlaGlyPheArgPheGlyAlaTyIleuLeuGlnAlaGl 885  
Db 1863 CTTTATATATTTTGGCCTATGATGAGCAGGGTTTCGATTTTGGAGCCTATTATTAATTCAGAGTCG 1922

QY 885 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyTyGlyAlaMetAl 905  
Db 1923 ACGAATGACCCAGAGGGCATGTTTCATAGCTTTTACTGCAATTCATATGGAGCTATGGC 1982  
QY 905 alleGlyLysThrLeuValLeuAlaProGluTySerLysAlaLysSerGlyAlaAlaHi 925  
Db 1983 CATCGAAAAACGCTCGTTTGGCTCTGAATATTCCAAAGCCAAATCGGGGCTCGCA 2042  
QY 925 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLy 945  
Db 2043 TCTGTTTGGCTTGTGGAAAAAGAACCAATATAGACACCGCAGTCACAGAGGGAATA 2102  
QY 945 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyTyProCysAr 965  
Db 2103 GCCAGACACATGTGAAGGAAATTTAGAGTTTCAGAAAGTCTCTTTCTCTATCATGTGC 2162  
QY 965 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 985  
Db 2163 CCAGATGTTTTCATCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222  
QY 985 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1005  
Db 2223 ATTTTGGGAGCAGCGCTGTGGAAAAAGACACTTCTGTTCAACTTCTGCAGAGACTT 2282  
QY 1005 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1025  
Db 2283 TGACCCCGTCAAGGACAAAGTGTGTTGATGATGATGCAAAAGAAATTTGAATGTACA 2342  
QY 1025 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1045  
Db 2343 GTGGCTCCGTTCCTCAATAGCAATCGTTCTCAAGAGCCCTGTCTCTTCAACTTCAGC 2402  
QY 1045 eAlaGluAsnIleAlaTyGlyAspAsnSerArgValValProLeuAspGluLysGl 1065  
Db 2403 TGTGTGAACATCGCTATGTGACAAACAGCCGTGGTGCCATTAGATGAGATCAAGA 2462  
QY 1065 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyTyAsnTh 1085  
Db 2463 AGCCGCAATGCAGCAAAATATCCATTTTATTTGAAGTCTCCCTGAGAAATACAAC 2522  
QY 1085 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1105  
Db 2523 ACAAGTTGAGCTGAAAGGAGCAGCTTTCTGGCGCCAGAAACAAAGACTAGCTATTGC 2582  
QY 1105 aArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAs 1125  
Db 2583 AAGGCTCTTCTCCAAAAACCCAAAAATTTTATTTGTTGGATGAGCCACTTCAGCCCTCGA 2642  
QY 1125 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1145  
Db 2643 TAATGACAGTGAAGAGTGGTTTCAGCATGCCCTTGATAAAGCCAGGACGAGGACATG 2702  
QY 1145 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1165  
Db 2703 CCTAGTGGTCACTACAGGCTCTCTGCAATTCAGAACGCGAGATTTGATAGTGTCTGCA 2762  
QY 1165 sAsnGlyLysIleLeuGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1185  
Db 2763 CAATGAAAGATTAAGAAACAGGAACTCATCAAGAGCTCTCGAGAAATTCGACATATA 2822  
QY 1185 rPheLysLeuValAsnAlaGlnSerValGln 1195  
Db 2823 TTTTAAAGTAGTAATGCACAGTCAGTCAGTCAG 2853

## RESULT 6

CRUPGP1

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

CRUPGP1 4296 bp mRNA linear ROD 22-JUN-1995  
C.griseus p-glycoprotein (isoform class I) mRNA, complete cds.

M60040

M60040.1 GI:191164

multidrug resistance; p-glycoprotein; transmembrane protein.

C.griseus adult liver and ovary, cDNA to mRNA.



[illegible]



Qy	418	AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg	437
Db	1549	GA AAAACATTCGCTATGGCCGAGAAAATGTCTATGGACGAGATCGAGAAAGCTGTCTCAAG	1608
Qy	438	GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly	457
Db	1609	GAAGCCAAATGCCCTATGACCTTCATCATGAAGATGCCCCATAAAATTGACACCTCTGTGTGGT	1668
Qy	458	GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu	477
Db	1669	GAGAGAGAGACAGCTGAGTGGAGACAGAAAACAGAGAAATCGCCATTCCTCTGTCGCCCTG	1728
Qy	478	ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer	497
Db	1729	GTCCGCAACCCCAAGATCCTTTTGTGGATGAGGCGACATCAGCCCTGGACACAGAAAGT	1788
Qy	498	LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal	517
Db	1789	GAAGCCGTGGTTCAGCGCTGCTCGGATAAGGCTAGAGAAAGCCCGACCTACCATTTGTGATA	1848
Qy	518	AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet	537
Db	1849	GCCACCGCTTGTCTACAGTTCGAAATGCTGCATCATCCTAGCTGGGTTTGATGGTGGTGC	1908
Qy	538	LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu	557
Db	1909	ATTGTGGAGCAAGCAAAATCATGAAGAGCTCATCAGAGAGAAGGCGATTACTTCAAACTT	1968
Qy	558	ValMetSerGln-----AspIleLysLysAlaAsp	567
Db	1969	GT CATGACACACAGCAGAGAAATGAAATTAGGAATGAAGTTGGTGAGCTATAA	2028
Qy	568	GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----Pro	585
Db	2029	AATGAAATTCATAATTTAGACATGCTCTCAAAGATTCCAGCATCCAGCTCTAATCAGAAGG	2088
Qy	586	LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluLysSerThr---	604
Db	2089	AGATCAACTCCGACAGATATCCGTGGACACACATGACCAAGACAGGAAGCTTAGTACCAAA	2148
Qy	605	GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLys	624
Db	2149	GAGCCCTTGGATGAGATGATCCTCCAAATTTCTTTTGGCGTATCCTAAAGATTGAATCA	2208
Qy	625	ProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHis	644
Db	2209	TCTGAATGGCCATATTTTGTGGTTGGTATATCTGTGCCATAGTAATGGAGCCTTGCAA	2268
Qy	645	ProValPheSerIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLys	663
Db	2269	CCAGCATCTCTCAATAATATTTCTAAGGTTGTAGGGGTTTTTCACAAGAAATACTGATGAT	2328
Qy	664	ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle	683
Db	2329	GAACCAAAACACATGATACCACTGTTTTTCTTATGTTTCTGATCTCTGGGGTCAAT	2388
Qy	684	CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr	703
Db	2389	TCTTTTATTACATTTTCTTCAGGCGCTTCACATTTGGCAAAAGCTGGAGAGATCCTCACC	2448
Qy	704	MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp	723
Db	2449	AAGCGATCCCGATACATGGTTTTTCAATTCATGCTGACAGAGATGTGAGCTGGTGTGAT	2508
Qy	724	GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle	743
Db	2509	AACCTTAAAAACACCATGGAGCAATTCACACAGCGCTCGCCACAGATCGCTGGTCAAGTT	2568
Qy	744	GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetClyLeu	763
Db	2569	AAAGGGGTACAGAGACCACTGTGTCTTTCATTTACCCAGAAATAGCAAACTATAGCAAACT	2628

Qy	764	Serv	Val	lle	lle	Ser	Ph	e	lle	Tyr	Gly	Trp	Gl	u	Met	Thr	Ph	e	Leu	lle	Leu	Ser	lle	ala	783	
Db	2629	GGA	TAT	CA	TAT	AT	CC	C	T	A	T	A	T	C	T	A	T	C	T	A	T	C	T	A	T	2688
Qy	784	Pro	Val	Leu	Ala	Val	Thr	Gly	Met	ile	Glu	Thr	Ala	Ala	Met	Thr	Gly	Ph	e	Ala	Asn	Lys		803		
Db	2689	CC	CA	T	CA	T	GC	A	T	A	T	GC	A	T	A	T	GC	A	T	A	T	GC	A	T	2748	
Qy	804	Asp	Lys	Gln	Leu	Lys	His	Ala	Gly	Lys	lle	Ala	Thr	Glu	Ala	Leu	Glu	Asn	ile	Arg				823		
Db	2749	GAT	AA	GA	AG	AG	CT	AG	AG	CT	CT	GG	AA	GT	T	CT	AA	GA	CA	AT	AG	AA	CT	2808		
Qy	824	Thr	ile	Val	Ser	Leu	Thr	Arg	Glu	Lys	Ala	Ph	e	Glu	Gln	Met	Tyr	Glu	U	Me	Leu	Gln		843		
Db	2809	ACT	GC	CT	CT	CT	TT	G	ACT	CGG	AG	C	AG	AA	GT	T	GA	AA	AT	AT	GT	AT	GC	2868		
Qy	844	Thr	Gln	His	Arg	Asn	Thr	Ser	Lys	Lys	Ala	Gln	lle	lle	Gly	Ser	Cys	Tyr	Ala	Ph	Ser		863			
Db	2869	AT	AC	CA	T	AC	CA	GA	AA	T	CT	CT	GA	AA	AG	CA	CA	CG	T	CT	TT	GG	AT	2928		
Qy	864	His	Ala	Ph	e	lle	Tyr	Ph	e	Ala	Tyr	Ala	Ala	Gly	Ph	e	Arg	Ph	e	Gly	Val	Tyr	Leu	lle	883	
Db	2929	C	AG	CC	AT	GA	T	AT	TT	T	CG	T	AT	CT	CT	CT	TT	CC	GG	TT	T	GG	CT	2988		
Qy	884	Ala	Gly	Arg	Met	Thr	Pro	Glu	Gly	Met	Ph	e	lle	Val	Ph	e	Thr	Ala	lle	Ala	Tyr	Gly	Vala	903		
Db	2989	CG	GA	AT	TA	AT	GA	CA	TT	G	AA	AT	GT	CT	AT	T	AT	T	AT	T	CT	C	AG	3048		
Qy	904	Met	Ala	lle	d	ly	Lys	Thr	Leu	Val	Leu	Ala	Pro	Glu	Tyr	Ser	Lys	Ala	Lys	Ser	Gly	Vala		923		
Db	3049	AT	GC	AG	CT	CGG	CA	GT	CA	GT	TC	CT	CA	T	AT	GC	CA	T	AT	GC	CA	AA	AG	3108		
Qy	924	Ala	His	Leu	Ph	e	Ala	Leu	Glu	Lys	Lys	Pro	Asn	lle	Asp	Ser	Arg	Ser	Gln	Glu	Gly		943			
Db	3109	T	CG	AT	AT	CA	T	AT	CA	T	AT	CA	T	AT	CA	T	AT	CA	T	AT	CA	AG	3168			
Qy	944	Lys	Lys	Pro	Asp	Thr	Cys	Glu	Gly	Asn	Leu	Glu	Ph	e	Arg	Glu	Val	Ser	Ph	e	Tyr	Pro		963		
Db	3169	CT	GA	AG	CT	TA	AT	CA	TT	T	GA	AG	AA	AT	GT	G	AA	AT	TA	AT	GA	AG	3228			
Qy	964	Cys	Arg	Pro	Asp	Val	Ph	e	lle	u	Arg	Gly	Leu	Ser	Leu	Ser	lle	Glu	Arg	Gly	Lys	Thr		983		
Db	3229	ACC	CG	CG	CT	CA	CA	CT	CC	AG	T	GC	TT	CA	GG	GG	CT	GA	AC	CT	GG	AG	3288			
Qy	984	Val	Ala	Ph	e	Val	Gly	Ser	Ser	Gly	Lys	Ser	Thr	Ser	Val	Gln	Leu	Gln	Arg				1003			
Db	3289	CT	GC	CC	T	GT	GG	CG	AG	CA	GT	GC	T	CG	GG	AA	AG	CA	CA	GT	GC	T	CG	3348		
Qy	1004	Leu	Tyr	Asp	Pro	Val	Gln	Gly	Gln	Val	Leu	Ph	e	Asp	Gly	Val	Asp	Ala	Lys	Glu	Leu	Asn		1023		
Db	3349	TT	CT	AC	CA	CC	CC	AT	GC	CG	GA	CA	AG	T	TT	CT	AG	AT	GC	CA	AA	AG	3408			
Qy																										

```

Db 3709 CTGGATACAGAGAGTGAAGAGTGTCCAGAGAGCTCTGGACAAAGCCAGAGAGGCCGC 3768
Qy 1144 ThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaLeuValVal 1163
Db 3769 ACTGCATTGTGATCTCACCCTGTCCACCATCCAGAACCCAGACTGTATGCTGGTG 3828
Qy 1164 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 1183
Db 3829 ATTCAGAAATGGCAAGTCAAGGAGCAGTGGCACCACCCAGCAGCTGTCTGGCACAGAAAGC 3888
Qy 1184 IleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3889 ATCTATTCTCCATGTCAGTGTGCAGGCT 3918

RESULT 7
CRUPGP1165 CRUPGP1165 4304 bp mRNA linear ROD 27-APR-1993
LOCUS Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION M59253
VERSION M59253.1 GI:191154
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.
SOURCE Chinese hamster, cDNA to mRNA, clone ADX165.
ORGANISM Cricetulus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
REFERENCE 1 (bases 1 to 4304)
AUTHORS Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
TITLE Full length and alternatively spliced pgpi transcripts in
multidrug-resistant Chinese hamster lung cells
JOURNAL J. Biol. Chem. 266 (7), 4545-4555 (1991)
MEDLINE 91154265
PUBMED 1671863
FEATURES
     source
       1..4304
         /organism="Cricetulus sp."
         /db_xref="taxon:10031"
         /cell_type="fibroblast"
         /tissue_type="lung"
       1..4304
         /gene="pgp-1"
         /db_xref="GI:191155"
         /translation="MEFEEDFSGRKDKNFKLGRKSKKEKPKPVSVFTMFRVAGW
         LDRYMLVGTAAIIGHVALPLMLVDFGMDTDSFASVGNIPNTNNAQVNASDIFG
         KLEEMTVAYYVTGTGAGLVIVAYVIONVFWCLAAAGROIHKROKFFHAIHQEITGF
         DVHVGELNTRLDVSKINEGDKIGMFFORMATFFGGFIIGTRGKHLTLVILAI
         SPVLGSLGKAWKILSSFTDKELQAYAKAGAAEVLAAIRTVIAFGGKKELRYNN
         NLEEKARLIGKILANTANISMGAAFLIYASALAFWYGLSVISKEYSIGQVLTFFFA
         VLIAPFSIQGASPEIEAFANARGAAEIEFNIIDNPSIDFSKNGYKPNIDNIGLEFK
         NTHFSPSKDQVILKLNKQSGOTVALVNGSGCKSTTVQLLQRLVDPTGVVSI
         DQDQRTINVRYLREITGVVSOEPLVFTATIAENIRYGRNVMTDEIEKAVEANAYD
         FIMKLPHKFTDLVGRGAQLSGQKORIAIALALVRLNPKILLIDNATSALDTESEAVV
         QALDKAREGRTIIVIAHRLSTIADIDFGVIVEQGNHELMREKGIYFKLVM
         TQAGNEIPLSEVEGSKNEIDNLMSSKDSASSLIRRRSTRSIRGPHDQRLKSTK
         EALDEDPVPSIFSRILKLNSEWPFVVGIFCAIVNGALQAPAFSIFSVKVGFRNT
         DDEIKRHSLSLLPLIGVLSFTFFLQGTFFGKAGELTKRLRYVMFKSMLRQDV
         SWFDNPKNTGALTFLANDAGVKGATGARLAVITONIANLGTGIIISLIYGMQLTL
         LLLAVPIIAAGVEMKMSQALKDKKELESGSKIATEIENRTVVSLSREQKE
         NMYAGSLQIPYRNALKHAFVGTFTFTQAMMIFSTAACFRFGALVARELMFENVL
         LVFSAIVFGAMQVQSSPAPYAKAKVSASHIIMIEKVPISIDSYSGGLKPNLELG
         NVKFEVFNFTRPDIPLVGLNLEVKKGOTLALVSSCGSKSTVQLLERFYDFMA
         GTVFKDGKVNOLNQLRHLGIVSOEPLFDCSAENIAVGDNSRVVSOEIERAA
         KEANTHOFESLPDKYNTRVGDKGTOLSGQKORIAIALALVROPHILHLDATSLD
         TESEKVVQALDKAREGRTIIVIAHRLSTIADIDFGVIVEQGNHELMREKGIYFKLVM
         IYFSWVSVQAGAR"
       1275 a 895 c 1031 g 1103 t
       Map position 1q26.

```

BASE COUNT  
ORIGIN

```

Alignment Scores:
Pred. No.: 3,04e-223 Length: 4304
Score: 3573.00 Matches: 679
Percent Similarity: 75.37% Conservative: 248
Best Local Similarity: 55.20% Mismatches: 259
Query Match: 59.43% Indels: 44
DB: 10 Gaps: 8

US-09-873-409-6 (1-1195) x CRUPGP1165 (1-4304)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 258 ATGCTGGTGGGACTCTCGCAGCTATTATCCATGGAGTTGCATCCCGCTTATGATGCTG 317
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn----- 38
Db 318 GTGTTTGGAGACATGACACAGCTTTGCAAGTGTAGGAAACATCCCAACAAATGCTACT 377
Qy 38 ----- 38
Db 378 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 437
Qy 39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db 438 ACGTACGCCCTAT-----TACTACACAGGATTTGGTCTGTGCTGTATATA 482
Qy 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db 483 GTTGCTTACATTGAGTTTCATTTTGGTGCCTGGCAGCTGGGAAGACAAATACACAAAT 542
Qy 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db 543 AGACAGAAGTTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGACGTGCAGC 602
Qy 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db 603 GTTGGGGAGCTCAACACCCGGCTCACAGATGATGCTCTCCAAATTAACGAAGAAATGGT 662
Qy 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 663 GACAAAATTTGGATGTTCTTCAGGCAATGGCAACAATTTTTGGTGGTTTATATAGGA 722
Qy 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 723 TTTACTCGAGGCTGGAAGCTAACCTTTGTGATTTTGGCCATCAGCCCTGTTCTTGAC 782
Qy 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyr 177
Db 783 TCAGCTGGTATTTGGGCAAGATATTATCTTCAATTTACTGATAAAGAACTTCAGGCATAT 842
Qy 178 SerIleAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 843 GCAAAAGCTGGAGCAGTCTGCTGAAGAGTCTTAGCAGCCATCAGAACTGGATTGCAAT 902
Qy 198 ArgAlaGlnGlyLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 903 GAGGACAAAGAAAGAACTTGAAGGTACACAACTATTAGAAGAGCCAAAGACTT 962
Qy 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db 963 GGGATAAAGAAAGCAGTACACAGCAACATTTCCATGGGCGCAGCTTCTCTTCTTATCTAT 1022
Qy 238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerIleLeuAsnGlyLeuProGly 257
Db 1023 GCATCATATGCTCTGGCTTCTGGTATGAAGACCTCTTGGTCTCTCTCTCTTATCTAT 1076
Qy 258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
Db 1077 TATCTATTGGACAAAGTGTCTACTGTCTTCTTCTTGTGTTATTAATTCGCCCATTCAGTATT 1136
Qy 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297

```





Qy	199	AlaGlnGluysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	218
Db	827	GGACAAAGAAAGAACTCTGAAGAGTACACAAAAATTTAGAGGAAGCTTAAAGGAATTTGGG	886
Qy	219	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	238
Db	887	ATAAAGAAGCTATACCGGCAACATTTCTATGTGGCGGCTTCTTATTCATTCATTCATGCA	946
Qy	239	ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	258
Db	947	TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCTCTCCAGTGAA-----TAT	1000
Qy	259	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	278
Db	1001	ACTATTGGACAAGTACTCACTGTCTCTTTCTGTATTAAATTTGGGGCTTTTAGTATTGGA	1060
Qy	279	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	298
Db	1061	CAGGCATCCCCAAGCAATTGAAGCATTTGCAACCGCAAGAGGAGAGCTTATGAATCTTC	1120
Qy	299	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	318
Db	1121	AGAATAATTGACATAAACCAGCATTCAGACCTATTGCAAGAGTGGACATAAACACGAT	1180
Qy	319	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	338
Db	1181	AATATTAAAGGAAATTTGGAAATTCAAAAATGTTCACTTCAGTTACCTCTTCGAAAAAGAA	1240
Qy	339	IleIysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
Db	1241	GTTAAGATCTTTAAAGGCTCTCAACCTGAAGGTTCCAGAGTGGCGACAGAGTGGCGCTGGT	1300
Qy	359	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
Db	1301	GGNAACAGTGGCTCGGGAGAGCACGCCGTGCAGCTGATGCGAGAGGCTCTATGACCCC	1360
Qy	379	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	398
Db	1361	ACAGATGGCATGCTGTGTATTCATGACAGGACATTAGGACCAATAATGTAAAGCATCTT	1420
Qy	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
Db	1421	CGGGAATTAATCTGGTGTGGTGAAGTCCAGAGCCCTGTGTGTTGCCACCACCATGAGCTGAA	1480
Qy	419	AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu	438
Db	1481	AACATTGCCTATGGCGCGGAAATGTCCACATGGATGAGATTGAGAAAGCTGTTAAGGAA	1540
Qy	439	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	458
Db	1541	GCCAAATGCCCTATGATTTTATCATGAATACTACCTAATAAATTTGACACTCTGTTGGAGAG	1600
Qy	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	478
Db	1601	AGAGGGGCCCGAGCTGAGTGGTGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGT	1660
Qy	479	ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
Db	1661	CGCAACCCCAAGATTTCTTCGTGGATAGGCAACGTGAGCTCTGGACATCGAAAGTGAA	1720
Qy	499	SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla	518
Db	1721	GCAGTGGTTCAGGTGGCCCTGGATAGGCCAGAAAGGCCGAGACTACCATTTGTGATAGCT	1780
Qy	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
Db	1781	CATCGTTTGTCTACAGTTCGTAATCCGATGTCAATGTGCTGGTTTTGTGATGGAGTCAAT	1840
Qy	539	AlaGluLysGlyAlaHisIaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	558
Db	1841	GTGGAGAAAGAAATCATGATGAACATCATGAAGAGGAGGCACTTACTTCAAACTTGTCT	1900
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578

Db	1901	ACATGCAG---ACAAGACGAATGGAATTGAGTTAGAAAATGCCACTGGTGATCCAAA	195
Qy	579	ArglysThrAsnSerLeuProLeuHisSerVallylsSerllelySerAspPheIle---	597
Db	1958	AGTGAAGATGATCCCTTGGAAATG--TCTCCAAAAGATTACAGGCTCCAGTTAATAAAA	2014
Qy	598	-----AspLysAlaGluGluSer	603
Db	2015	AGAAGATCAACTCGCAGGAGTATACATGCACCAAGGCCAAGACAGAAAGCTTGGTACA	2074
Qy	604	ThrGlnSerlysgLulleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
Db	2075	AAAGAGACTTGAATGAGAATGTACTCCAGTTTCCTCTCGAGATTCTGAAGCTGAAC	2134
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2135	TCAACTGAATGGCTTATTTTGTGGTGGTATATTTTGTCTATTATAAACGGAGGCGTG	2194
Qy	644	HisProValPheSerIlelePheAlaLysIleleThrMetPhe---GlyAsnAsnAsp	662
Db	2195	CAACGACATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT	2254
Qy	663	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2255	CCTGAACAACAAACGACAGAATAGTAACATGTTTCTGTATTGTTTCTAGTCCCTTGGAAAT	2314
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2315	ATTTCCTTTATACATTTTCTCCAGGCTTCACATTTGGCAAAAGCTGGGAGATCCTC	2374
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2375	ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGACAGCAGGATGTCAGCTGGTTT	2434
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2435	GATGACCTTAANAACCACTCGAGCATTTGACCAACCGGCTTGCCAATGATGCGGCTCAA	2494
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2495	GTTAAAGGGCTATAGTTCAGGCTTGCTGTCATATCCAGAAATATAGCAAACTTTGGG	2554
Qy	763	LeuSerValIleleSerPheIleTyrGlyTrpGluMetThrPheLeuLeuLeuSerIle	782
Db	2555	ACAGGCATTTATATCCTTAATCTATGTTGGCAATTAACACATTTTACTCTTAGCAATT	2614
Qy	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2615	GTACCCCATCAATGCAATAGCAGAGTTGTGAAATGAAAAATGTTGTCGGAACAAGCACTG	2674
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
Db	2675	AAAGATAAGAAAGAGCTAGAGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAATCTC	2734
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2735	CGAACTGTGTTCTTGACTCGGAGCAGAAGTTTGAATACATGATGCACAGAGTTTG	2794
Qy	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	862
Db	2795	CAAGTACCATACAGAAACTCTTTGAGGAAGACACATCTTCGGGGTCTCTATTTTCTATC	2854
Qy	863	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	882
Db	2855	ACCCAGGCAATGATGATATTTTCTATGCTGGCTGTTTCCGGTTTGGTGCTACTTGGTG	2914
Qy	883	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleArgly	902
Db	2915	GCAATGAGTTCATGAACITTCAGGATGTTCTTTGGTATTCTCAGCATTTGCTCTTGGT	2974
Qy	903	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaIleArgSerGly	922

Db	2975	GCATGCGAGTGGGCGAGTCACTTTCCTGCTCACTATCCCAAGCAAGTATCA	3034
Qy	923	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	942
Db	3035	GCAGCCACGTCATCATGATCATTTGAAAGAGCCCTCTGATTGACAGCTACAGCCCTCAC	3094
Qy	943	GlyLysLysProAspThrCysGluGluGluAsnLeuGluPheArgGluValSerPhePheTyr	962
Db	3095	GGCCTCAAGCCAAATACGTTGGGAAGGAAATGTGACATTTTAATGAGTGTGTTCAACTAT	3154
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
Db	3155	CCCACTGACACGACATCCCGTGCTCCAGGGCTGAGCCTCGAGTGGAAGAGGCCAC	3214
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	1002
Db	3215	ACGCTGCCCTCGTAGTAGCAGTGGCTGTGGGAAGACACAGCTTGTTCAGCTCCTAGAG	3274
Qy	1003	ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3275	CGCTTCATGACCCCTTGGCTGTGCTCAGTGCTGCGAAGACACAGCTTGTTCAGCTCCTAGAG	3334
Qy	1023	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3335	AATGTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCATCTGTTTGAC	3394
Qy	1043	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu	1062
Db	3395	TGCAGCATTCGCCAGAACATTCGCTATGGAGACACACAGCGGGTCGTATCACATGAAGAG	3454
Qy	1063	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	1082
Db	3455	ATTGTCAGGCGAGCAAGGAGGCGCAACATACACACACTTCATCGAGACACTCCCTCGAGAA	3514
Qy	1083	TyrAsnThrClnValGlyLeuLysGlyAlaGlnLeuSerClyGlyGlnLysGlnArgLeu	1102
Db	3515	TACAAACAGCAGTAGAGAGCAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT	3574
Qy	1103	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	1122
Db	3575	GCATAGCTCGCGCTTGTGTAGACAGCCTCATATTTGCTTTGGATGAAGTACATCA	3634
Qy	1123	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	1142
Db	3635	GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTCGACAAAGCCAGAGAAGGC	3694
Qy	1143	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	1162
Db	3695	CGCACCTGCTATGTCGCGCCACCGCTTGTCCACCATCCAGAAATGCGAATTAATAGTG	3754
Qy	1163	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	1182
Db	3755	GTGTTTCAGATGGCAAGTCAAGAGCATGGCCACATCCACATCCAGCTGTGCGCCAGAA	3814
Qy	1183	AspIleTyrPheLysLeuValAsnAlaGlnSer	1193
Db	3815	GGCATCTATTTTCCATGGTCACTGTCCAGGCT	3847
RESULT	9		
AX108656			
LOCUS			
DEFINITION	Sequence 3 from Patent WO0123565.		
ACCESSION	AX108656		
VERSION	AX108656.1		
KEYWORDS	crab-eating macaque.		
SOURCE	Macaca fascicularis		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;		
	Cercopitheciinae; Macaca.		
REFERENCE	1 (bases 1 to 4195)		
AUTHORS	Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.I.		
TITLE	P-glycoproteins from macaca fascicularis and uses thereof		
Patent:	WO 0123565-A 3 05-APR-2001;		
GENTEST	CORPORATION (US)		
Location/Qualifiers	1..4195		
FEATURES			
source	/organism="Macaca fascicularis"		
	/db_xref="taxon:9541"		
CDS	100..3951		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC37765.1"		
	/db_xref="GI:13923889"		
	/translation="MDLEGRNGGAEKKNFKLNKSKKDKKERTVSVFMSFRYSN		
	WLDKLYMVVGLAALIHGAGLPLMLVFGDMTDTFANAGNLGLDGLALLPNTNNSNIT		
	DTVPVNLNEDMTYAYIYSGIGAGLVAAIYQVSWFCLAGRQIHKIRKQPFPHAIR		
	QEIQFVDVHVGELNTRLTDVSGINEGIDKIMGFQSMATFFFTFVIGTFRGKMLT		
	LVLIAISPLGLSAAWAKILSFTDKELLAYAKAVAEVLAIAIRTVIAFGQKKE		
	LERYNKLLEAKRIGIKKAITANISIGAFILLIYASVALAPWYGTITLVLSKYSIGQV		
	LTVPSVLIGAFSYGQASPSIEAFANARGAFAEIKIIDNKPSIDYSKSGHKPNIK		
	GNLFRNVHPSYPRKREVKILKGLNLKQSGQTVALVNSGCGKSTTVQLMQRLDYPT		
	EGMSVDGQDITINVRFLREIIGVVSQEPVLFIATINIRYGRHEDVTMDIEKAVK		
	ENAYDFIMKLPQKFDITIVGERQLSGGQKQRIARALVRNPKILLIDLEATSLDT		
	ESEAIVQALDKARKGRTTIVIAHRLSTVRNADVIAGFDGVIKESGHNDELMEKGI		
	YKLTVMOTAGNEILENADESKEIDTLEMSHSDSGSLIRKESRRSVRSQSQD		
	RKLSKEALDESIIPVSWRIMKLNLTMPYVGVFCALINGLOPAFAVFSKILIG		
	IFTRNDRAETKRONSLFLFLVLGIVSFITFLOGGTFPGKAGBILKRLRYMWFERS		
	MLRQDVSFDDPKNTTGTALTRLANDAAQVKAIGSLRLAITQNTIANLGTGIIISLY		
	QMQLTLLLLAIPIIAIAGVVMKMLSQGLKDKKELEGAKIATIEINFTVVSLLT		
	QEQFPHMYDOSLVQVYRNSLRKAHIFGITFTFQAMMYFVAGCFRFGAYLVASHLM		
	SFEDVLVFSVAGMAMVGVSSPAPYAKAKVSAHIIMTIEKTPIDISYSTGLK		
	PNTLEGNVTNEVFNYPTRLDIPVLQSLSEVKKGTALVLVSGSGCKSTVQLIER		
	FYDPLAGKLLDGGREIKQNLVONWRAHIGIVSQEPILFDCSISENIAIYQNSRVVSQE		
	EIVRAAKNTHAFIESLPNKYSTRVGDKTQSGQKQRIARALVRNPKILLIDLE		
	ATSALDTSEKVVQALDKAREGRTTIVIAHRLSTQNLADLIIVFQNGRVEKHGTHQ		
	LLAOKGIVFSVMVQAGAKRO"		
BASE COUNT	1230 a 801 c 1039 g 1125 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,49e-222	Length:	4195
Score:	3556.50	Matches:	680
Percent Similarity:	75.16%	Conservative:	246
Best Local Similarity:	55.13%	Mismatches:	263
Query Match:	59.16%	Indels:	43
DB:	6	Gaps:	9
US-09-873-409-6 (1-1195) x AX108656 (1-4195)			
Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	250	ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG	309
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSer-	33
Db	310	GTGTTTGGAGACATGACGGATACCTTTGGCAATGACGAGAAATTTAGGAGATTAGGAGCT	369
Qy	34	LeuValGlnThrAsnThrTyrSer-	41
Db	370	CTGTTCTTTAAACAACAATAATAGCAGTAAATATCACTGATACAGTCCCGTCATGAATCTG	429
Qy	42	-----PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValalalaLeuIle	58
Db	430	GAGGAAGATATACACAGGATGCTTATTTATACAGTGGAAATTTGGTGGTGGCTGGT	489
Qy	59	PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIle	78
Db	490	GCTGCTTACATTCAGGTTTCATTTTGGTGGCAGCTGGGAAGACAAATACACAAAT	549
Qy	79	ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp	98
Db	550	AGAAACAGTTTTTTCATGCTATAATCGCACAGGAGATAGGCTGGTTTGTATGTCAGCAT	609
Qy	99	IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly	117



Db 610 GTTGGGAGCTTAACACCCGGCTTACAGATGATGTCTCAAGATTAAATGAAGAAATTGGT 669  
Qy 118 AsplysilealalaLeuPheGlnAsnMetSerThrPheSerileGlyLeuAlaValGly 137  
Db 670 GACAAAATTGGAATGTTCTTTCAGTCATGCGCAACATTTTTCATCTGGTTTATAGTAGGA 729  
Qy 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuLeuMetAla 157  
Db 730 TTTACAGTGGTGGAGACTAACCCCTGTGATTTTGGCCATCAGTCTCTGTTCTGGACTG 789  
Qy 158 SerAlaAlaCysSerArgMetValLysSerLeuThrSerLysGluLeuSerAlaTyr 177  
Db 790 TCAGCTGCACTGGGCAAGATACCTCTTCATTACTGATAAGAACTCTTAGCTTAT 849  
Qy 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerileArgThrValLysAlaPhe 197  
Db 850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGCTTGTGAGCAATTAGAACTGTGATGCAATT 909  
Qy 198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217  
Db 910 GGAGGACAAAAGAAAGAACTCGAAAGGTCAACACAAAATTTAGAAAGACTAAAGAAATT 969  
Qy 218 GlyLysArgThrLysAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237  
Db 970 GGGATAAAGAAAGCTATTACAGCCCAATATTCTATAGGTGCTCTTTCCTGCTTATCTAT 1029  
Qy 238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAenGlyGluProGly 257  
Db 1030 GCATCTTATGCTCTGGCCCTCTGTGATGGGACCACTTGTCTCTCAAGGAA----- 1083  
Qy 258 TyrThrileGlyThrValLeuAlaValPhePheSerValLysSerTyrCysile 277  
Db 1084 TATCTATTGGACAAGTACTCACCTGTTCTTCTGTTAAATTTGGGCTTTTAGTGT 1143  
Qy 278 GlyAlaAlaValProHisPheGluThrPheAlaAlaArgGlyAlaAlaPheHisile 297  
Db 1144 GCACAGGCATCTCCAAAGCATTGAAGCATTTCGCAATGCAAGAGAGCAGCTTTTGAATC 1203  
Qy 298 PheGlnValLysLeuAspLysProSerileAspAsnPheSerThrAlaGlyTyrLysPro 317  
Db 1204 TTCAGATAATTGATAATAAGCCAAAGTATTGACAGCTATTGCAAGAGTGGGCAACCA 1263  
Qy 318 GluSerileGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337  
Db 1264 GATAAATTAAAGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAA 1323  
Qy 338 SerLysileLeuLysGlyLeuAsnLeuArgLysSerGlyGluThrValAlaLeu 357  
Db 1324 GAAGTTAAGATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTG 1383  
Qy 358 ValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377  
Db 1384 GTTGGAAACAGCGCTGTGGGAAGAGCACACCGTCCAGCTGATGACAGCGCTTTATGAC 1443  
Qy 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397  
Db 1444 CCCACAGGGCATGGTCAGTGTGATGAGCAGCAGATATTAGCACCAATAAGAGTGT 1503  
Qy 398 TyrArgAspHisileGlyValValSerGlnGluProValLeuPheGlyThrThrileSer 417  
Db 1504 CTACGGGAAATCATCGGTGCTGAGTCAGGCAACCTGTATTGTTTGGCCACCAATAGCT 1563  
Qy 418 AsnAsnileLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437  
Db 1564 GAAACATTCGCTATGGTCGTGAGAGATGTCACCATGATGAGATTGAGAAGCTGTCAAG 1623  
Qy 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457  
Db 1624 GAAGCCATGCTATGACTTATCATGAACCTGCCTCAGAAATTTGACACCTCGTTTGA 1683  
Qy 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgLysGlnArgAlaLeu 477  
Db 1684 CAGAGGGGGCCAGCTGAGTGGTGGGCGAGAGAGAGTGGCCATTGGCAGCTGGCCCTG 1743

Qy 478 ValArgAsnProLysileLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497  
Db 1744 GTTCGCAACCCCAAGATCTCTCTGTGACGAGGCCACGTCAGCCTTGGACACAGAACT 1803  
Qy 498 LysSerAlaValGluAlaAlaLeuGluLysAlaSerLysGlyArgThrThrileValVal 517  
Db 1804 GAAGCAGTGTGTCAGGTGCTCTGATTAAGGCCAGAAAAGGTGCGGACCACTTGTGATA 1863  
Qy 518 AlaHisArgLeuSerThrileArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537  
Db 1864 GCTCATCGTTTGTCTACGGTTCGTAAATGCCAGCGCATCGCTGGTTCGATGATGGAGTC 1923  
Qy 538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557  
Db 1924 ATTTGGAGAAAAGAAATCATCATGAGCTCATGAAAGAGAAAGGCATTTACTTCAACTT 1983  
Qy 558 ValMetSerGlnAspIle-----LysLysAlaAspGlu----- 568  
Db 1984 GTCACAATCGACACAGCAGGAAATGAAATTTGAATTAGAAAATGCGAGCTGATGAATCCAAA 2043  
Qy 569 ---GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu----- 584  
Db 2044 AGTGAATTTGATACCTTGGAAATGCTTCACATGATTCAGGATCCAGTCTAATAAGAAAA 2103  
Qy 585 -----ProLeuHisSerValLysSerileLysSerAspPheIleAspLysAlaGluGlu 602  
Db 2104 AGATCCACTCGTAGGAGTGTCCGTGGATCAAGGCCAA-----GACAGAAAGCTTAGT 2157  
Qy 603 SerThrGlnSerLysGluLysSerLeuProGluValSerLeuLeuLysileLeuLysLeu 622  
Db 2158 ACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTATGAAGTGA 2217  
Qy 623 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 642  
Db 2218 AATTAACTGAGTGGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2277  
Qy 643 ValHisProValPheSerilePheAlaLysileIleThrMetPhe---GlyAsnAsn 661  
Db 2278 CTGCAACACGACATTCGACGTAATATTTTCAAAGATTATAGGGATTTTACAAGAAATGAT 2337  
Qy 662 AspLysThrThrLeuLysHisAspAlaGluLysSerMetLysPheValLysLeuGly 681  
Db 2338 GATGCGCAAAACAAAACGACAGAAATAGTAATCTTTTCACTATTGTTCTTAGTCTTGA 2397  
Qy 682 ValLysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluile 701  
Db 2398 ATGTTTCTTTTATATACATTTTCTTCAGGCTTCACATTTGGCAAGCTGGAGAGATC 2457  
Qy 702 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyr 721  
Db 2458 CTCACCAAGCGGCTCCGATACATGTTTCCGATCCATGCTCAGACAGGATGTGAGCTGG 2517  
Qy 722 PheAspGluLysGluAsnSerThrGlyLeuThrThrileLeuAlaIleAspIleAla 741  
Db 2518 TTTGATGACCTTAAAAACACCACTGGAGCATTGACTACAGGCTCGCCCAATGATGTGCT 2577  
Qy 742 GlnleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 761  
Db 2578 CAAGTTAAAGGGCTATAGTTCAGGCTTGCTATTAATTTACCAGAAATATAGCAATCTT 2637  
Qy 762 GlyLeuSerValLysileSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSer 781  
Db 2638 GGGACAGGAATAATTATCTTAAATCTATGTTGGCAACTGACACTGTACTCTTAGCA 2697  
Qy 782 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 801  
Db 2698 ATTTGATCCCATCTATGCAATAGCAGGAGTGTGTAATGAAATGTTGTTGTGACCAAGCA 2757  
Qy 802 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysileAlaThrGluAlaLeuGluAsn 821  
Db 2758 CTGAAAGATGAAGAACTAGAAAGTGGGAGAGATCGCTACTGAGCAATAGAAAAC 2817



QY 822 IleArgThrIleValSerLeuThrArgGluIlyeAlaPheGluGlnMetTyrGluGluMet 841  
Db 2818 TTCGAACTGTTGTTCTTTGACTCAGCAGCAGAGTGTGAACATATGATGATCAGAGT 2877  
QY 842 LeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 861  
Db 2878 TTCGAGTACCATACAGAACTCTTTCAGGAAAGCACAATCTTTGGAAATCAGTTTCC 2937  
QY 862 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 881  
Db 2938 TTCACGCGAGCAATGATGATATTTTCCTATGCTGGATGTTTCGGTTGGAGCCTACTTG 2997  
QY 882 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 901  
Db 2998 GTGCACATAGTCTCAGAGCTTTGAGGATGTTCTGTGTATTTTCAGCTGTGCTTT 3057  
QY 902 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 921  
Db 3058 GTTGCCATGGCGTGGGCAAGTCAGTTCATTGCTCTGACTATGCAAGCCAAAGTA 3117  
QY 922 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGln 941  
Db 3118 TCAGCAGCCACATCATCATGATCATTTGAAATAAACCCCTTTGATTGACAGCTACAGCAC 3177  
QY 942 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 961  
Db 3178 GAAGGCGTAAAGCGCAACACATTTGGAAGGAATGTCACATTTAATGAAGTTGATTCAAC 3237  
QY 962 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 981  
Db 3238 TATCCCAACCGACTGGACATCCCGAGTCTTCAGGGGCTGAGCCCTGGAAGTGAAGAAGGCG 3297  
QY 982 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 1001  
Db 3298 CAGACGCTGCGCTGTTGGGACAGTGGCTGTGGGAAGCAGCGGTGTCAGCTCTG 3357  
QY 1002 GlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlu 1021  
Db 3358 GAGCGGTTCATGACCCCTTTGGCGGGAAAGTGTGCTTGACGCCAAAGAAATAAGCAA 3417  
QY 1022 LeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 1041  
Db 3418 CTGAATGTTTCAGTGGCTCCGACACACCTGGCATGCTGTGCCAGGAGCCCATCTCTGTT 3477  
QY 1042 AsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAsp 1061  
Db 3478 GACTGCAGCATTAGTGAGACATTCCTATGGAGACACAGCCCGGTGTGTACAGGAA 3537  
QY 1062 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 1081  
Db 3538 GAGATCGTGAGGCGAGCAAGGAGGCCAATATACACGCTTCATCGAGTCACTGCCTAAT 3597  
QY 1082 LysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 1101  
Db 3598 AATATAGCACACAGTAGGAGCAAGAACTCAGCTCTCTGTGTGCCAAGAACACAGC 3657  
QY 1102 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 1121  
Db 3658 ATTGCCATGCTGTCCTCTGTTAGACGCTCATATTTGCTTTTGGAGTGAAGCCACA 3717  
QY 1122 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1141  
Db 3718 TCAGCTCTGGATACAGAAAGTGAAGAGTTGTCCAAAGCCCTTGGCAAGCCAGAGAA 3777  
QY 1142 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1161  
Db 3778 GGGCGTACCTGCAATGTGATTTGCTTCCACCGCTGTCCACCATCCAGATGCGAGACTTAATA 3837  
QY 1162 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsn 1181  
Db 3838 GTGGTGTTCAGATGGCAGATCAAGGACAGCGACATCAGCAGTGCCTGCGACAG 3897  
QY 1182 ArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193

Db 3898 AAAGGCATCATTTTCAATGGTCAAGTGTCCAGGCT 3933  
RESULT 10  
AF257746  
LOCUS  
DEFINITION  
Rattus norvegicus multidrug resistance protein 1a (Pgly1) mRNA.  
complete cds.  
AF257746  
AF257746.1 GI:7739772  
Rattus norvegicus  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 4927)  
Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,  
Meijer,D.K.F. and Muller,M.  
Cloning and functional characterization of the rat multidrug  
resistance protein Mdr1a  
Unpublished  
REFERENCE  
2 (bases 1 to 4927)  
Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,  
Meijer,D.K.F. and Muller,M.  
Direct Submission  
Submitted (19-APR-2000) Div. Gastroenterology & Hepatology,  
University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,  
The Netherlands  
FEATURES  
Location/Qualifiers  
1..4927  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/sex="male"  
/cell\_type="hepatocyte"  
/tissue\_type="liver"  
/dev\_stage="adult"  
1..4927  
/gene="Pgly1"  
86..3904  
/gene="Pgly1"  
/note="ATP-binding cassette transport protein; Abcb1a"  
/product="multidrug resistance protein 1a"  
/protein\_id="AAF69007.1"  
/db\_xref="GI:7739773"  
/translation="MELEDNLGRADKNFPMKSKSKSKKKPAVSULTMFRYAGW  
LDRFVLLGLTAAIIGHGIALPLMLVFGDMTDSFANVGNRSMSPFNATDIYAKUEDE  
MTTYAYTGTGAGVLIYAIQVSLWCLAAGRIHKIRQKFFHAINRQDIWGFVHDV  
GELNRLTDDVSKINEGIDKIMPFQAMATPFGGFIIGFTGWLTLVILAI SPVLG  
LSAGTWAKTISSTDKELQAYAKAGAAVEVLAAITVIAFGQKELERYNNLEEA  
KELGIKKAITANISGAAPFLIYASYALAPFWGTSLVTSKEYTIQVLTVPFSLIGA  
FSVQASPNIEAFANAGAAVEFSIIDNKPSIDFSGSKHDPNIOGMLERNIHFS  
YPSRNVQVILKGLNLKVSQGTVALVNSGCGKSTVQLLQRLYDPIEGEVIDGDI  
TRINDYLRLEIIGVVSQBPVLPATTIAENIRYGRNVTMDIEKAVKEANAYDF INKL  
PHKPTLVGERGAQLSGGOKORIAIARALVRNPKILLDEATSAIDTSEAEVVOAALD  
KAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHDELMREKGIYFKLVMTQTAG  
NEIEGNACESKSGIDINDMSKSGSLIRRRSTRKSRPHQDQGLSLTKEALDD  
DVPPASFWRILKLNSTWPFYVGVFCALINGLOPAFSIIFSVKVSFVKNDTPIQ  
RQNSNLSLLFLILGIIIPITPFLQGFPGKAGEILTAKRLRYMVKSMRLQDI SWPFD  
PKNTTGALTTRLNDAAGVKATGSRNAVITONIANLGTGIIISLYGQLTLILAI  
VPILAIAGVEMKMLSGOALKDKLEGSGKIATBAIENFRVTVSLTREQKPTMYAQ  
SLQIYRNALKAHVFGITFTFOAMYPFSAACFRFGAYLVARELMTENVLLVPSA  
IVGFAMVQVSSFPADYAKAVKVSASHIIRIEKIPEIDSYSTEGKPNLNSGNVFN  
GVFNYPTRPNIPVLOGLSLEVGSGCGKSTVVOLLERFYDPMAGTVFL  
DKETIKQLNVWLRLHGLTVSQEPILFDCSTAEANIAYGNSRVSHEEIVTKAEANI  
HOFIDSLPEKYNTRVHDKGTGSLGQKORIAIARALVRNPKILLDEATSAIDTSEK  
VVOEALDKAREGRTTIVIAHRLSTIQNADLIIVIQNGQVKHGTHTQQLLAQGIYFSM  
VSVQAGAKRS" 1474 a 1037 c 1190 g 1226 t  
BASE COUNT  
ORIGIN

## Alignment Scores:

Pred. No.: 4,98e-222 Length: 4927  
Score: 3555.50 Matches: 682  
Percent Similarity: 75.78% Conservative: 244  
Best Local Similarity: 55.81% Mismatches: 263  
Query Match: 59.14% Indels: 33  
DB: 10 Gaps: 9

US-09-873-409-6 (1-1195) x AF257746 (1-4927)

```
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysIleuProLeuMetProLeu 20
DB 233 ATGCTGCTGGAACTCTGGCGGCAATATCATGTGAATTCGCTCCCACTTATGATGCTG 292
QY 21 ValIleuGlyGluMetSerAspAsnLeuIleSerGlyCysIleuValGlnThrAsnThrTyr 40
DB 293 GTCTTTGGAGACATGACAGATAGCTTTGGCAAAAT-----GTAGGAACAACCGTAGTATG 346
QY 41 SerPhePheArgLeuThr----- 46
DB 347 AGTTTCTACAACTCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCAGTACGCC 406
QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
DB 407 TACTATTACACGGGCATTTGTCGGGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 466
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgIysGlnPhePheHisSerVal 86
DB 467 TGTGCTCGCGAGCTGGAGACAAATACACAAGATTAGGACGAAGTTTTTCCATGCCATC 526
QY 87 LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
DB 527 ATGAATCAGAGATAGCTGCTGGTTGACGTGCATGACGCTTGGGAGCTCAACACCGGCTC 586
QY 107 Thr----AspIleAspIysIleSerAspGlyIleGlyAspIysIleAlaLeuLeuPheGln 125
DB 587 ACAGATGACGCTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTTCAG 646
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIysGlyTyrIleLeuThr 145
DB 647 GCAATGGCAACATTTTGTGGTGTATATATAGGATTTTACTCGCGCTGGAAGCTAACT 706
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
DB 707 CTGTGTGATTTGGCCATCAGCCCTGTTCTTGGACTGTGCTGATTTGGGCAAGATA 766
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
DB 767 TTGTCTTCATTTACTGATAAGGAATCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 826
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
DB 827 GAAGTCTTAGCAGCCATCAGAACTGTGATTTGCCCTTTGGAGGACAAAAGAACTTGAA 886
QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
DB 887 AGGTACAAATAACAAATTTGGAAAGAGCTTAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 946
QY 226 LysValSerLeuGlyAlaValTyrPheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245
DB 947 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATATGCTCTGGCACTTCG 1006
QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
DB 1007 TATGGGACTTCTTGGTCACTCTCAAAAGAA-----TACACTATTGGACAAGTGTCTCACT 1060
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
DB 1061 GTCTTTTTCGTATTAAATGGAGCAATTCAGTCTGGCGAGCATCTCCAAATATTGAA 1120
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 305
DB 1121 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAATAATGATAAATAAGGCC 1180
```

```
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
DB 1181 AGTATAGACAGCTTCTCAAGAGTGGGCACAAACCGACACATACAGGAAATTTGGAA 1240
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
DB 1241 TTCAAAATATTTCACTTCACTTACCGTCTCGAAAAGAGCTTCAGATCTTTGAAGGCGCTC 1300
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
DB 1301 AACCTGAAGGTGAAGAGCGGCGAGCGTAGCCCTGTTGGCAACAGTGGCTCTGGGAAA 1360
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
DB 1361 AGCACAACTGTCACAGCTGCTGCAGAGGCTCTACGACCCCTAGAGGGCGAGGTCAGTATC 1420
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
DB 1421 GACGGACAGACATCAGGACCATCAATGTGAGTATCTCGGGAATCATTTGGGTGGTG 1480
QY 406 SerGlnGluProValIleuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
DB 1481 AGTCAGAAACCCGCTGCTGTTGCCACCAATTTGCCGAAACATTCGCTATGCGCGAGAA 1540
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
DB 1541 AACCTCACCATGATGATAGAGAAAGTGTCAAGAGAGCCATGCTATGATTTCAATC 1600
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyGlyAlaGlnMetSerGly 465
DB 1601 ATGAAACTCCCCACAAATTTTGACACCTTGTGGTGAGAGGGCGGCGAGTGGG 1660
QY 466 GlyGlnGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
DB 1661 GGACAGAAACAGAGGATCGCATTTGCCCGGCCCTGCTCGCAACCCCAAGATCTTTTG 1720
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
DB 1721 TTGATAGAGCCACGTCAGCCTTGGACACAGAAAGCGAGCGTGGTTCAGGCGGCTCTG 1780
QY 506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
DB 1781 GATAAGGCTAGAGAGCGGCGGACCACTTGTAGAGTCTCACCGCTTGTCTACAGATTCG 1840
QY 526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
DB 1841 AATGCTGACGCTCATCTGCTGCTTGTATGTTGTGGTGTCTATTGTGGAGCAAGGAATCATGAT 1900
QY 546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln-----MetThrTyr 561
DB 1901 GAGCTCATGAGAGAAAGGAATTTACTCAAACTTGTCTATGACTCAGACAGCAGGAAAT 1960
QY 562 AspIleLysLysAlaAspGluGlnMetGluSer-----MetThrTyr 575
DB 1961 GAAATTTAGTAATAGGAATGAAGCTTGTGAATCTAAAGATGGAATGTAATGGGACATG 2020
QY 576 SerThrGluArgLysThrAsnSerLeu-----ProLeuHisSerValLysSerIleLys 593
DB 2021 TCTTCAAAAGATTTCAGGATCCAGTCTCTAATAAGAAAGAGATCAACTCGCAAAAGCATCCGT 2080
QY 594 SerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeuPro 612
DB 2081 GGGCCCATGATCAAGCGGGGAACCTAGCACCAAGAGGCTCTGGATGACGACGTACCT 2140
QY 613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyrProPheValValLeu 632
DB 2141 CCAGCTTCTTTTGGCGGATCTCTGAAGTTGAATTCAACTGAATGGCTTATTTTTGGTT 2200
QY 633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
DB 2201 GGTGTATTTTGTGCCATAATAAATGGAGGCTTCGAACCCAGCATCTCTCCATAATATTTTCA 2260
```

```
QY 653 LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAlaGlu 671
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2261 AAGTTGTAGGGGTTTTTACAAAATATGACACCCCTGAAATCCAGCGGAGAACAGCAAC 2320

QY 672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2321 TTGTTTCTTATGTTCCGTGATCCCTGGATCATCTCTTTCAATTACGTTTTTCCCTTCAG 2380

QY 692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2381 GGCTTCACATTTGGCAAGCTGGAGAGATCCTCAACCAAGCGACTCCGATACATGTCCTTC 2440

QY 712 LysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGluAsnSerThrGlyGly 731
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2441 AATCCATGCTGAGACAGACATAAGCTGGTTGATGATGACCCCTAAACACACACAGAGCG 2500

QY 732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyValThrGlySerArgIle 751
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2501 CTGACCACAGGCTTGCCAAATGACGCTCAAGTGAAGGGGCTACAGGCTCTAGGCTT 2560

QY 752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2561 GCTGTTATTACCAGAACATAGCAATCTTGGGACAGGCATCATATCCCTGATCTAC 2620

QY 772 GlyTyrGluMetThrPheIleLeuLeuSerIleAlaProValLeuAlaValThrGlyMet 791
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2621 GGCTGGCAATTGACACTTTTACTCTAGCAATTTGTTCCCATCATCTGATAGCAGAGTG 2680

QY 792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2681 GTTGAATGAAATGTTGCTGCAACAGGCTGAAAGATAAGAAAGAACTAGAAAGTTCT 2740

QY 812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2741 GGGAAATCGCTACAGACGAATGAAACTTTCCGACGTGCTCTCTTTGATCCGGAG 2800

QY 832 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 851
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2801 CAGAAGTTTGAAACTATGATGCCAGAGGCTTGCAATACCAATGCTTTGGAAG 2860

QY 852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2861 AAGCCGACGCTTTGGGATCATCTTCTCCTCCAGCCGCAATGATGATTTCTCCTAT 2920

QY 872 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 891
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2921 GCTGCTGTTCCGTTTGGTGTGCTACTGCTGGCGACGAGAACTCATGATTTGAAAT 2980

QY 892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2981 GTTCTGTAGTATTCAGCTATTGTTGTTGGTGCCATGGCGAGGTCAGTTCA 3040

QY 912 LeuAlaProGluTyrSerLysAlaLysSerGlyValAlaAlaHisIleLeuPheAlaLeuGlu 931
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3041 TTGCGCTCCTGACTACGGAAGCAAGTCTCGGCATCCACATCATCATGATCATTTGAG 3100

QY 932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3101 AATATCCCTGAGATTGACAGCTACAGCAGGAGGCTTGAGGCTATATGTTGGAGGA 3160

QY 952 AsnLeuGluPheArgGluValSerPheThrProCysArgProAspValPheIleLeu 971
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3161 AATGTGAATTTAATGGAGTCATGTTCAACTATCCACCCGACCAACATCCCAAGTCTT 3220

QY 972 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3221 CAGGGGCTGAGCCTAGAGTGAAGAAAGGCGACAGCGTGGCCCTCGTGGCGACAGTGGC 3280

QY 992 CysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAspProValGlnGlyGln 1011
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3281 TGGCGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGTCGGCAACA 3340

QY 1012 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1031
```

```
Db 3341 GTGTTCTAGATGGCAAAATAAGCAACTCAATGTCCAGTGGCTCGCGCCACCTG 3400
QY 1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3401 GGCATGTGTCCAGAGGCCATCTGTTGACTGCAGATCGCCGAGAACATTTGCTTAC 3460

QY 1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAsn 1071
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3461 CGAGACAAACAGCCGTGCTGTCATGAGGAGATCGTGAAGCGACGCAAGGAGGCAAC 3520

QY 1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3521 ATCCACCAAGTTTCATCGACTCACTGCTGAGAAATCAACACAGAGTGGGAGCAAAAGG 3580

QY 1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuGlnLys 1111
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3581 ACTCAGTGTGCGGGGGGAGAGAGCGCATCGCATCGCGGGCCCTCGTCAGACAG 3640

QY 1112 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3641 CCTCATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGAGAGTGAAGAGTGC 3700

QY 1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3701 GTCCAGGAAGCGCTGGACAAAGCCAGGAGCGCACCTGTCATGTCGCGCACCGC 3760

QY 1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1171
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3761 CTGTCCACCATCCAGAACCGACATTGATCTGTGTGATTCAGAACCGCCAGGTCAGGAG 3820

QY 1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3821 CACGCGACCCACAGCAGCTGCTGGCCCAAGAAAGCATCTATTCTCGATGCTCAGTGTG 3880

QY 1192 GlnSer 1193
|||:|||||:
Db 3881 CAGGCT 3886

RESULT 11
LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192877.
ACCESSION AX322787
VERSION AX322787.1 GI:18093766
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sorrentino, B. and Schuetz, J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 1 06-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
source
1. 3860
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1135 a 746 c 957 g 1022 t
ORIGIN
Alignment Scores:
Pred. No.: 4,56e-222 Length: 3860
Score: / 3554.00 Matches: 682
Percent Similarity: 75.63% Conservative: 246
Best Local Similarity: 55.58% Mismatches: 263
Query Match: 59.12% Indels: 36
DB: 6 Gaps: 9
US-09-873-409-6 (1-1195) x AX322787 (1-3860)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
```

Db 151 ATGGTGGTGGAACTTTGGCTGCATCATCCATGGGGCTGACTTCCTCTCATGATGCTG 210  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 211 GTGTTGGAGAAATGACAGATATCTTTGCAATGCAGGAAATTTAGAAATCTGATGTC 270  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 271 AACATCACTAATAGAACTGATATCAATGATACAGGGTTCCTCATGAATCTGGAGGAAGAC 330  
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 331 ATGACCAGATATGCTATTATTACAGTGAATTTGGTCTGGGTGCTGCTGTAC 390  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleValAlaAlaLeuIleGlyGln 81  
Db 391 ATTCAGGTTTCATTTTGGTCCCTGGCAGCTGGAGCAAAATACACAAATTTAGAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGCTATATGACACAGGATAGGCTGGTTTGTATGTGCACGATTTGGGGAG 510  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspIleVal 120  
Db 511 CTTAAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTATTGGTGCACAAAT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGTCTCTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GGTGGAGCTAACCTTGTTGATTTGGCCATCAGTCTCTGTTCTTGACACTGTCAGCTGCT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 691 GTCTGGCAAGATACTATCTTACTTACTGATAAAGAACTCTTAGCGTATGCCAAAGCT 750  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGAGCAGTACTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 810  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLysAspAlaLysAspPheGlyIleLys 220  
Db 811 AAGAAAGAACTTGAAGGTACACAAATTTAGAAAGAACTTAAAGAAATTTGGATAAG 870  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 871 AAGCTATTACAGCAATATTCTTATAGTGTCTGCTTTCTCTGCTGATCTATGATCTTAT 930  
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 931 GCTCTGGCTTCTGATGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 984  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 985 GGACAAGTACTCACTGATTTCTTCTGATTAATTTGGGCTTTTGTGTTGACAGGCA 1044  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1045 TCTCCAAGCATTTGAAGCATTTTGAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1104  
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1105 ATTGATAATAGCCCAAGTATTGACAGCTATTTCGAAGAGTGGCGCAACACAGATAATT 1164  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1165 AAGGGAATTTGGAAATTCAGAAATTTCACTTCACTTACCATCTCGAAAGAGTTAAG 1224  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360

Db 1225 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGCAGACCGTGGCCCTGTTGGAAAC 1284  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1285 ACTGGCTGTGGGAGAGCACAACAGTCCAGCTGATGCAGAGCTCTATGACCCACAGAG 1344  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1345 GGGATGGTCACTGTTGATGCAGAGGATATTAGGACCATATAATGTAAGGTTTCTACGGGA 1404  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420  
Db 1405 ATCATTTGGTGTGGTGAACCTGTATTGTTGCCACCACGATAGTGAACAAT 1464  
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1465 CGCTATGGCGTGAATATGTCACCATGGATGAGATTGAGAAAGCTCTCAAGNAGCCAT 1524  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
Db 1525 GCTATGACTTTTATCATGAACCTGCTCATAAATTTGACACCTGTTGGAGAGAGGG 1584  
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1585 GCCCAGTTCAGTGGTGGGAGAGCAGAGGATCGCCATTTGCACGTGCCCTGGTTCGCAAC 1644  
Qy 481 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1645 CCCAGATCTCTCTGCTGGTGGAGCCACGTGAGCTTGGACACAGAAAGCAGAGCTG 1704  
Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAlaHisArg 520  
Db 1705 GTTCAGTGGCTCTGATAAGGCCAGAAAGGTTCGACACCATTTGTGATAGTCTATCGT 1764  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 1765 TTGTTACAGTTCGTATGATGTCACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
Db 1825 AAGGAAATCATGATCACTCATGAAGAGAGAAAGGCAATTTACTTCAAACTTGTCAATG 1884  
Qy 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGACGATGAATCCAAAGTGAAT 1944  
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
Db 1945 GATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGTCCAGTCTTAATAAGAAAGATCAACT 2004  
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607  
Db 2005 CGTAGAGTGTGCTGGATCAGAGCCAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
Qy 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627  
Db 2065 GATGAAAGTATACCTCCAGTTCTTCTTTGGAGGATATTGAAGCTAAATTTAACTGAATGG 2124  
Qy 628 PropheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
Db 2125 CTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2184  
Qy 648 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 2185 GCAATATATTTTCAAGATATTAGGGTTTTTACAGAAATTCATGATCTCTGAAACAAA 2244  
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
Db 2245 CGACAGATAGTAATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2304  
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706  
Db 2305 ACATTTTCTTCAAGTTTTCACATTTTGGCAAGCTGGAGATCTCTCACCAGCGGCTC 2364

```
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyxGlnAspIleAlaTrpPheAspGluLysGlu 726
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2365 CGATACATGTTTCCGATCCATGCTCAGACAGCATGTGAGTTGGTTGATGACCCCTAAA 2424

QY 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2425 AACACCACTGGACATGTGATACAGGCTCGCAATGCTGCTCAAGTTAAAGGGGCT 2484

QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2485 ATAGGTTCCAGGCTGTGTAATATACCCAGATATAGCAATCTTGGGACAGGAATAATT 2544

QY 767 IleSerPheIleTyArgIleGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2545 ATATCTCTCATCTATGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCAT 2604

QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2605 GCATAGCAGGAGTGTGTGAATGAAATGTTGTCGGACACAGCACTGAAAGATAAGAAA 2664

QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2665 GAACTAGAGGCTCGGAAGATCGTACTGAAGCAATAGAAACTTCCGAAACGGTTGTT 2724

QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrgluGluMetLeuGlnThrGlnHis 846
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2725 TCCTTGACTCAGGACGAGAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyralaPheSerHisAlaPhe 866
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2785 AGAACTCTTTGAGGAAAGCACACATCTTTGGAAATTACATTTTCTTCACCCAGGCAATG 2844

QY 867 IleTyxPheAlaTyxAlaAlaGlyPheArgPheGlyAlaTyxLeuIleGlnAlaGlyArg 886
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2845 ATGTAATTTTCCATGCTGATGTTCCGGTTTGGACCTACTTGTGGGCACATAAATC 2904

QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyxGlyAlaMetAlaIle 906
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2905 ATGAGCTTTCAGGATGTTCTGTAGTATTTCAGCTGTTGCTGTTGGTCCATGGCCGTG 2964

QY 907 GlyLysThrLeuValLeuAlaProGluTyxSerLysAlaLysSerGlyAlaAlaHisLeu 926
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2965 GGCAAGTCAGTTCATTTGCTCTGATATGCCAAGCCAAATATACAGAGCCCAATC 3024

QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3025 ATCATGATCATGAAAACCCCTTTGATTGACAGCTACAGCAGGAGCCCTAATGCCG 3084

QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyxProCysArgPro 966
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3085 AACACATTTGAAGGAATGTCACATTTGGTGAAGTGTATTCAACTATCCACCCGACCG 3144

QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3145 GACATCCAGCTGCTTCAGGAGCTGAGCTCGAGGTGAAGAGGCCAGAGCGCTGGCTCTG 3204

QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyxAsp 1006
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3205 GTGGCAGCAGTGGCTGTGGAGAGACACAGTGTCTCAGCTCTCTGGAGCGGTCTACGAC 3264

QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3265 CCCTTGGCAGGAAGTGTCTGTATGGCAAGAAATAAAGCACTGATGTTTCAGTGG 3324

QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3325 CTCGAGACACATCGGTCATCGTGTCCAGGAGCCCATCTCTGTTTGCATCGACATTCGT 3384

QY 1047 GluAsnIleAlaTyxGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3385 GAGAAACATTCCCTATGTGAGACAAACAGCCGGTGTGTGCAGGAAGAGATCTGTGAGGCA 3444
```

```
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyxAsnThrGln 1086
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3445 GCAAGAGAGGCCCAACATACATCTCTCATCGATCTACTGCTAATAATATATACACTAAA 3504

QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnArgLeuAlaIleAlaTyx 1106
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3505 GTAGGAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACACGCAATGGCCATGCTCGT 3564

QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3565 GCCTTGTAGACAGCCTCATATTTTGTCTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3625 GAAAGTGAAGAGGTTGTCGAAGAACCCCTGGCAAAAGCCAGAGAGCCGACCTGCAT 3684

QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3685 GTGATTGCTCACCGCTCTCCACCATCCAGATTCAGACTTAATAGTGTGTTTCAGAA 3744

QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyxPhe 1186
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3745 GGCAGAGTCAAGGAGCATGCGCAGCATCAGCAGCTCTGCGCACAGAAAGCATCTATTT 3804

QY 1187 LysLeuValAsnAlaGlnSer 1193
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3805 TCAATGTCAGTGTCCAGGCT 3825

RESULT 12
AX105078 AX105078 4279 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 22 from Patent WO0123540.
DEFINITION AX105078
ACCESSION AX105078
VERSION AX105078.1 GI:13921228
KEYWORDS
SOURCE dog:
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
Patten,C.J.
P-glycoproteins and uses thereof
Patent: WO 0123540-A 22 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
CDS 17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37721.1"
/db_xref="GI:13921229"
/translation="MDPEGGRKGSAAEKNFMKMGKSKKNEKKEKKTSTVTFAMFRYSN
WLDRLYLVTGMAAIIHGAALPLMLVFNMTDSFANAGISRNKTFPVIINISNTNT
OHFINLHEEMTVAYYVYEGVAGVVAAYIOVSFWCLAAGROIILKROKFOHAIKRL
ETGWFVDHVDGELNTRLTDDVSKINEGIDKIMGFQSIATPFTGKIVGTGKWLTL
VILASPVLGSAAIWAKLSPTDKELLAYAKAGAAVEVLAARTVIAFGQKKEL
ERYKNLEBEAGIGIKKAITANISGAAPFLIIYASIALAFWTGTSILVSSEYSGQVL
TVFVSLVGFASIGQASPIEFAPNARGAAVEIFKIIDNKPSIDYSKSGHSPDNKIG
NLEFKNVHFSYPSRKEVILKGLNLKVSQGTVALVNGSCGKSTVQLMQRLYDPTD
GMVICDQDITRTINVRHLEITGVVQSPLVATTAEINRYGRNVTMDTEKAVKE
ANAYDFIMKLPNKFTLNGERGAOLSGGOKRIATARALVRNPKILLDEATSKDTE
SEAVQVALDKARKGRTTIVTAHRLSTRVADVIAGFDGDKVIRKGNHDELAKEGILY
FKLVTMQTRGNEIELENATGESDALEMSPKDSGSLIKRTRRSRIHAPQQODR
KLGTEDLENENVPVPSFWRILKLNSTWPFYFVVGFCATINGGLQPAFSIIISRIIGI
FTRDBDPTKRNQSNMFSLVLGIIISFITPFLQGTFGKAGELITKELRMYFRSM
LRQDYSWFDPPKNTGALTTLANDAAOVKAIGSLAVITONIAENLGTGIIISLIYR
WOLTLILLAIIVPIIAIGVEMKMLSGOALKKKELEGAKITAEIENFRVTVSLTR
EOKFEMYAQSLOVRYNRLKAHIFGVSFSTQMMYPSYAGCFRFGAYLVANBPMN
FODVLVFSFAIVFGNAYVQVSSFPADYAKAVKVAHVIMIIEKSPIDISYSPHGLKP
NTLEGNVTFNEVFNYPTRDPIPVILQGLSLEVKGQTLALVGSSGCGKSTVVQLLERF
```

YDPLAGSVLIDGKEIKHLNWMVLAHLGIIVSQEPILFDGSIENAIYAGDNRVWSHEE  
IMAAEKENHPIETLPKYNTRVGDKGTOLSGGOKRIAIARALVROPHILLIDEA  
TSLADTESKVVQEALEDKAREGRCTCVIAHRLSTIQNALDIIVFQNGKVEHGTQHL  
LAQGIYFSVMSVQAQAKR

BASE COUNT 1295 a 833 c 1008 g 1143 t  
ORIGIN

## Alignment Scores:

Pred. No.: 5 21e-222 Length: 4279  
Score: 3554.00 Matches: 682  
Percent Similarity: 74.74% Conservative: 238  
Best Local Similarity: 55.40% Mismatches: 269  
Query Match: 59.12% Indels: 42  
DB: 6 Gaps: 8

US-09-873-409-6 (1-1195) x AX105078 (1-4279)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 167 ATGTTGGTGGGACATGGCTGCCATCATCCATGGAGCTGCACCTCCCTCTCATGATGCTG 226  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGACAGAAATTTCAAGAAACAAAACCTTTT 286  
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 287 CCAGTTAATTAATGAAGTATTACGAACATACACAACTTTCATCAACCATCTGGAG 346  
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePhe 59  
Db 347 GAGGAATGACCATGATGCTATTATTACAGTGGATCGGTGGTGGCTGGCT 406  
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
Db 407 GCTTATCATCCAGGTTTCTTCTGCTGCTGCGCAGCAGAGACTACTCAAAATTAGA 466  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
Db 467 AACAAATTTTTTCAATGCTATCCGACAGGAGATGGCTGTTGACGTGCATGACGTT 526  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 527 GGGGAGCTTAACACCGGCTCACAGACGATGCTCCAAATCAATGAAGAAATTTGGCGAC 586  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 587 AAAATTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646  
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 647 ACAGTGGTTGGAAGCTAACCTTTGTGATTTTGGCCATCAGCCCTGTCTTGGACTTTCA 706  
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 707 GCGCCCATCTGGGCAAGATATCTTCTTCAATTTACTGATAAAGAACTCTTGGCCATGCA 766  
Qy 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 767 AAAGCTGGAGCAGTAGCTGAAGAGTCTTAGCAGCAATTCAGAACTGTGATGCGCTTTGGA 826  
Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
Db 827 GGACAAAGAAAGAACTTGAAGGTACACAAATTTTAGAAGACTTAAGAAATTTGGG 886  
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238  
Db 887 ATAAAGAAAGCTATCACGGCCCAACATTTCTATTGTCGCGCTTTCTTATTGATCATGCA 946  
Qy 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuLeuAsnGlyGluProGlyTyr 258  
Db 947 TCATATGCTCTGGCTTTCTGTTATGGGACCTCTTGGTCTCTCCAGTGAA-----TAT 1000

Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278  
Db 1001 TCTATTTGGCAAGTAGTACTACTGCTCTCTTTCTGTATTAATTTGGGCTTTTAGTATGGA 1060  
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaAargGlyAlaAlaPheHisIlePhe 298  
Db 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAAAACGCAAGAGGAGCAGCTTATGAATCTTTC 1120  
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
Db 1121 AGATTAATTCACATTAACCAAGCATTTGACAGCTATTGCAAGAGTGACATAACACAGAT 1180  
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
Db 1181 AATATTAAAGGAAATTTGGAATTTCAAAATTTTCACTTCAGTTACCTTCTCGAAAGAA 1240  
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358  
Db 1241 GTTAAGATCTTAAGGGTCTCAACCTGAAGTTTCAGAGTGGCAGACAGTGGCGCTGGTT 1300  
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378  
Db 1301 GGGACAGTGGCTGCGGAGAGCAGCAGCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360  
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398  
Db 1361 ACAGATGGCATGCTGTTATTCATGGACAGGACATTTAGGACCATTAATGTAAGGCATCTT 1420  
Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418  
Db 1421 CGGGAATTACTGGTGGTGGTGCAGGAGCCTGTGTTGTTGCCACCAGTAGCTGAA 1480  
Qy 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaAargGlu 438  
Db 1481 AACATTCGCTATGCGCGCAAAATGTCACCATGATGAGATTCAGAAAGCTGTTAAGAA 1540  
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458  
Db 1541 GCCAATGCCTATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTGGTGGAGAG 1600  
Qy 459 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaAlaLeuVal 478  
Db 1601 AGAGGGCGCCAGCTGATGCTGGACAAACAGAAATCGCCATCTCGGGCCCTGGTT 1660  
Qy 479 ArgAsnProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498  
Db 1661 CGAACCCCAAGATTTCTTCTGATGAGGACACGCTCAGCTCTGACACTGGAATGAA 1720  
Qy 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518  
Db 1721 GCAGTGGTTTCAGTGGCCCTGGATAAGGCCAGAAAGGCCGAGCTACCATTTGTATAGCT 1780  
Qy 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538  
Db 1781 CATCGTTGTCTACAGTTCGTAATGCCGATGTCATTTGCTGGTTTTCATGATGAGTCAAT 1840  
Qy 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558  
Db 1841 GTGGAAGAAAGAAATCATGATGAACTCATGAAAGAGAGGCGCATTTACTTCAACTGTC 1900  
Qy 559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578  
Db 1901 ACAATGCAG---ACAAGAGGAATGAATGAGTTAGAAAATGCCACTGGTGAATCCAA 1957  
Qy 579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597  
Db 1958 AGTGAAGTAGTCCCTTGGAAATG---TCTCCAAAAGATTACGGGTCAGGTTTAATAAAA 2014  
Qy 598 -----AspLysAlaGluGluSer 603  
Db 2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAGGCCAAGCAAGAGCTTGGTACA 2074  
Qy 604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 623



```
Db 2075 AAGAGGACTTGTAATGAGAAATGACCTCCAGTTTCCTCTGGAGGATTCGAAGCTGAAC 2134
QY 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db 2135 TCAACTGAATGGCTTATTTTGTGGTGGTATATTTGTGCTATTATAACGAGGACCTG 2194
QY 644 HisProValPheSerLeilePheAlaLysleileThrMetPhe--GlyAsnAsnAsp 662
Db 2195 CAACGACGATTTTCAATAATATTTTCAAGGATATAGGATCTTTACCCGAGATGAGAT 2254
QY 663 LysThrThrLeuLysHisAspAlaGluileTyrSerMetilePheValleileGlyVal 682
Db 2255 CCTGAACAAACACAGAAATAGTAAATGTTTCTGTATTTGTATTTCTAGTCCTTGAAT 2314
QY 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluileLeu 702
Db 2315 ATTTCTTTTATACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCCIC 2374
QY 703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetileuTyrGlnAspileAlaTyrPhe 722
Db 2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGACAGCAGGATGTCAGCTGTT 2434
QY 723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspileAlaGln 742
Db 2435 GATGACCCCTAAAAACACCACTGGAGCATTCGACCAACGAGGCTTCCTCAATGATGCGGCTCA 2494
QY 743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db 2495 GTTAAAGGGCTATAGTTTCCAGCTTCTGCTCATTTACCCAGAAATATAGCAATCTTGGG 2554
QY 763 LeuSerValileileSerPheileTyrGlyTyrGluMetThrPheLeuileLeuSerile 782
Db 2555 ACAGGCATTATATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614
QY 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db 2615 GTACCCATCATTCGAATGACGAGGATGTTTGAATGAAATGTTGTCTGACCAAGCACTG 2674
QY 803 LysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db 2675 AAGATGAAGAAAGCTAGAGAGCTGGGAAGATGCTACAGAAGCCATCGAAAACTTC 2734
QY 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2735 CGAACTGTTCTTTGACTCGGGACGAGAAGTTGAATACATGATGATGACAGAGTTTG 2794
QY 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleileGlySerCysTyrAlaPhe 862
Db 2795 CAAGTACCACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854
QY 863 SerHisAlaPheileTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuile 882
Db 2855 ACCGAGCAATGATGATTTTCTATGCTGGCTGTTTCCGGTTTGGTCTACTTGGTG 2914
QY 883 GlnAlaGlyArgMetThrProGluGlyMetPheileValPheThrAlaIleAlaTyrGly 902
Db 2915 GCAATGAGTTTCATGAATCTTCAGATGTTCTTTGTGATCTTCAGCTATGTTCTTGGT 2974
QY 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 2975 GCCATGCGAGTGGGCGAGGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAAGTATCA 3034
QY 923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db 3035 GCAGCCACGCTCATGATCATTTGAAATGAAAAAGCCCTCTGATGACAGCTACAGCCCTCAC 3094
QY 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db 3095 GGCTCTAGCCCAATATGTTGGAAGGAAATGTGACATTTAATAGGTCGTGTTCAACTAT 3154
QY 963 ProCysArgProAspValPheileLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
```

```
Db 3155 CCCACTCGACACAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAGGCCAG 3214
QY 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db 3215 ACGTGGCCCTCGTAGGTAGCAGTGGCTGTGGAAAGACACAGTGTCTCAGCTCTCCTAGAG 3274
QY 1003 ArgLeuTyrAspProValGlnGlnGlyValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3275 CGCTTCTATGACCCCTTGCTGCTTTCAGTGAATTTGATGGCAAGAGATAAAGCACCTG 3334
QY 1023 AsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGlnProValLeuPheAsn 1042
Db 3335 AATGTCAGTGGCTCGACACACCTGGGCATCGTGTCTCAGAGGCCCATCTCTGTTTGAC 3394
QY 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3395 TGCAGCATTTGCCGAGAACATTTGCTATGAGAGCAACACAGCCGGTCTGATCACATGAAGAG 3454
QY 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheileGluGlyLeuProGluLys 1082
Db 3455 ATTATGACGGCAGCCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAA 3514
QY 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3515 TACAACCCAGAGTAGGACAGAAAGAACCCAGCTCTCTGTGGCCAGAAACAGGCCATT 3574
QY 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db 3575 GCATAGCTGGGCTCTTCTGTAGACAGCTCATATTTTGTGTTTGGATGAAGCTACATCA 3634
QY 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3635 GCTCTGGATACAGAAAGTGAAGAGTGTCCAAGAGCCCTGGACAAAGCCAGAGAGGC 3694
QY 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspileuIleVal 1162
Db 3695 CGACCTGTGATGTGATCGCCACCGCTTGTCCACCATCCAGAAATGCAGATTTAATAGTG 3754
QY 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db 3755 GTGTTTCAGATGGCAAGTCAAGGACATGGCACACATCACAGCTGCTGGCCCGAGAA 3814
QY 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3815 GGCATCTATTTTCCATGTGTCAGTGTCCAGGCT 3847
RESULT 13
AX105080 LOCUS 4279 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 24 from Patent WO0123540.
ACCESSION AX105080
VERSION AX105080.1 GI:13921230
KEYWORDS dog.
SOURCE Canis familiaris
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 4279)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
Patten,C.J.
TITLE P-glycoproteins and uses thereof
JOURNAL Patent: WO 0123540-A 24 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
source Location/Qualifiers
1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37722.1"
/db_xref="GI:13921231"
/translation="MDPEGGRKGSAAEKFNKMGKSKKKEKKEKPTVSTFAMFRYSN
```





Qy	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	558
Db	1841	GTGAGAAAGGAAATCATGATGAATCATATGAAGAGAGAGGCACTTTACTTCACAACTGTGC	1900
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
Db	1901	ACAATGCGAG---ACAAGAGAAATGAAATTTGATTTAGAAAATGCCATCGTGGTAATCCAA	1957
Qy	579	ArgLysThrAenSerLeuProLeuHisSerValLysSerLleLysSerAspPheLle---	597
Db	1958	AGTGAAAGCTGATGCCCTGGAAATG---TCTCCAAAGGATTTCAGGGTCCAGTTTAATAAAA	2014
Qy	598	-----	603
Db	2015	AGAAAGATCAACTCGCAGGAGTATACATGCACCAAGCGCCAAGACAGAGAACTTGGTATCA	2074
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysLleLleLysLeuAsn	623
Db	2075	AAAGAGACTTGAATGAGAAATGACCTTCAGTTTCTCTCTGGAGATTCGTAAGCTGAAC	2134
Qy	624	LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2135	TCAACTGAATGGCTTATTTTGTGGTGTGTATATTTTGTGCTATTATTAACCGGAGCGCTG	2194
Qy	644	HisProValPheSerLleLlePheAlaLysLleLleThrMetPhe---GlyAsnAsnAsp	662
Db	2195	CNACCAGCATTTTCAATAATATATTTTCAAGGATATATAGGATCTCTTATCCCGAGATGAGAT	2254
Qy	663	LysThrThrLeuLysHisAspAlaGluLleTyrSerMetLlePheValLleLeuGlyVal	682
Db	2255	CCTGAAACAAACGACAGAAATAGTACATGTTTTCTGTATTGTTTCTAGTCTCTTGGAAAT	2314
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLleLeu	702
Db	2315	ATTTCTTTTATTACATTTTCTCCAGGCGCTTCACATTTTGGCAAGCTGGGAGATCCTC	2374
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPhe	722
Db	2375	ACTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTGAGCAGGATGTCCAGCTGGTGT	2434
Qy	723	AspGluLysGluAenSerThrGlyGlyLeuThrThrLleLeuAlaLleAspIleAlaGln	742
Db	2435	GATGACCTTAAACACCACTGGAGCATTTGACACCCAGGCTTGCCATGATCGGCTCAA	2494
Qy	743	IleGlnGlyAlaThrGlySerArgLleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2495	GTTAAAGGGGCTATAGCTTCCAGGCTTGCTGTGCTATTACCAGAAATATAGCAAACTTTGGG	2554
Qy	763	LeuSerValLleLleSerPheLleTyrGlyTyrGluMetThrPheLeuLleLeuSerLle	782
Db	2555	ACAGGCAATATATATCCTTAATCTATGTTGGCAATTAACACTTTTACTCTCTTAGCAAT	2614
Qy	783	AlaProValLeuAlaValThrGlyMetLleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2615	GTACCCATCATTCGATAGCAGAGTTGTTGAAATGAAAAATGTTGCTCTGGACAGCACTG	2674
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysLleAlaThrGluAlaLeuGluAsnLle	822
Db	2675	AAAGATAAGAAAGAGCTAGAAAGAGAGCTGGGAAAGATTGCTTACAGAGCCATCGAAAACTTC	2734
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2735	CGAACTGTGTTCTTTGATCTCGGAGCAGAGTTTGAATACATGATGTGCACAGAGTTTG	2794
Qy	843	GlnThrGlnHisArgAenThrSerLysLysAlaGlnLleLleGlySerCysTyrAlaPhe	862
Db	2795	CAAGTACCATACAGAACTCTTTGAGGAAGGCACACATCTTCCGGGCTCTCATTTTCTATC	2854
Qy	863	SerHisAlaPheLleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuLle	882
Db	2855	ACCCAGGCAATGATGATATTTTCTATGCTGGCTGTTTCCGCTTTGGTGCTACTGTGTT	2914

Qy	883	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyGly	902
Db	2915	GCAAATCAGTTTCATGAACCTTTTCAGGATGTTCTTTGGTATTTCTTCAGCTATTGTTCTTTGGT	2974
Qy	903	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTySerLysAlaLysSerGly	922
Db	2975	GCCATGTCAGTGGGACGTCAGTTCATTTGCTCTCTGACTATGCGCAAGGCCAAGATATCA	3034
Qy	923	AlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSerArgSerGlnGln	942
Db	3035	GCACCCACGTCATCATGATCATTTGAATAAGCCCTCTGATTGACAGCTACAGCCCTCAC	3094
Qy	943	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTy	962
Db	3095	GGCCTCAAGCCAAATACGTTTGAAGGAAATGTACATTTAATAGGTCGTGTCAACTAT	3154
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
Db	3155	CCCACCTCGACACACATCCCGCTGCCAGGGGCTGAGGCTCGAGGTGAAGAAGGGCCAG	3214
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	1002
Db	3215	ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGAGAGACACAGTTGTTTCAGCTCTTAGAG	3274
Qy	1003	ArgLeuTyrrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3275	CGCTTCTATGACCCCTTGGCTGGTTCAGTGTCTAAATTGATGGCAAGAGATAAAGCACCTG	3334
Qy	1023	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3335	AATGTCAGTGGCTCGAGACACCTTGGGCATCGTGTCTCAGAGGCCCATCTCTGTTGAC	3394
Qy	1043	CysSerIleAlaGluAsnIleAlaTyrrGlyAspAsnSerArgValValProLeuAspGlu	1062
Db	3395	TGCAGCATTCGCCAGAACATTCGCTATGAGAGACACAGCCGGGTCTATCATCATGAAGAG	3454
Qy	1063	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	1082
Db	3455	ATTATGCGAGCAGCAAGGAGGCCAACATACACCACTTCATCGACACACTCCCTGAGAAA	3514
Qy	1083	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	1102
Db	3515	TACAACACAGATGAGAGACAAAGGAACCCACAGCTCTCTGGTGGCCAGAAAACAGCGCAT	3574
Qy	1103	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	1122
Db	3575	GCCATAGCTCGCGCTCTGTGATGACAGCTCATATTTTCTTTGGATGAGCTACATCA	3634
Qy	1123	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	1142
Db	3635	GCTCTGGATACAGAAAGTGAAGAGTTGTCCAAGAGCCCTCGACAAAGCCAGAGAAAGGC	3694
Qy	1143	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspIleVal	1162
Db	3695	CGCACCTGCATTTGTGATCGCCACCGCTGTCCACCATCCAGAAATGCGAGATTAAATAGTG	3754
Qy	1163	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg	1182
Db	3755	GTGTTTCAGAAATGCAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGCCCCAGAAA	3814
Qy	1183	AspIleTyrrPheLysLeuValAsnAlaGlnSer	1193
Db	3815	GGCATCTATTTTCCATGGTTCAGTGTCCAGGCT	3847
RESULT 14			
LOCUS AX391099			
DEFINITION Sequence 1 from Patent WO0210205.			
ACCESSION AX391099			
KEYWORDS AX391099.1 GI:19584228			
SOURCE human.			
ORGANISM Homo sapiens			
linear DNA 4646 bp PAT 19-MAR-			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Ruth, A. and Roninson, I.

Mutations of the mdrl p-glycoprotein that improve its ability to

confer resistance to chemotherapeutic drugs

Patent: WO 0210205-A 1 07-FEB-2002;

THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

## FEATURES

Location/Qualifiers

1..4646

/organism="Homo sapiens"

/db\_xref="taxon:9606"

425..4267

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD28599.1"

/db\_xref="GI:19584229"

/translation="MDLEGRNGGAKKKNFFKLNKSEKDKKEKPTSVFVMFRYSN

WLDKLVVVGTLAAIIFHAGLPLMLVFGEMTDIFANAGNLEDLMSNITNRSIDINDTG

FMNLDEEDTRVAYVYGGIAGVLVAAVYQVSWCLAAAGROIHKIKQFFHAIMRQEI

GWFDVHDVGEINATRLDDVSKNEGIGDKIGWFOFOSMATFFTGFTVGRGKWLTLVI

LALSPVIGLSAANWAKILASPTDKELLAYAKAGAAVEALAIRVIAFGGOKKLEL

YKNLEAKRIGIKAITANI SIGAFLILYIAYALAFWITGLVLSGSEISYQGLVTV

FFSVLIGAFSVGQSPSIEAFANARGAAVEIFKIIDNKPSIDSYSKSHKPNIKGNL

EFNVHFSVSRKEVILKGLNLKVGQSTVALVNGSGCKSTVQLMQLDYDTEGM

VSDGQDITRTINVRFLREIIGVVSQEPVLPAITAEINRYGRNVTMDIEKAVKAN

AYDPIMKLPHKEDTLVGERGAOLSGGOKORIAALALVNPVKILLDEATSE

AVVQVALDKARKGRTIIVIAHRLSTVRNADVAGDDGVIVKGNHDELMKEGIFYK

LVTMTAGNELENADESKSEIDALEMSNDSRLIKRSTKSRVSGQAKRKL

STKEALDESIPVPSFWRIMKLNLTWPYFVWGVFCALINGGIQAPALIFSKLIGFT

RIDDPETKQNSNLSFLFLALGIIISFTFLQGTFFGKAGELTKRLRYVMVFRMLR

QDVSWFDKPNKLTALLTLANDAAOVKAIGSRLAVITONIANLGTGIIISFIYGMQ

LTLLALLVPIIAIAGVEMKMLSGOALKDKEGAKIATEAENRPTVLSLTQEO

KEHMYAQSLQPYRNSLKAHIFGTFSTQVMYPSVAGCFEGALVAHKLSPF

DVLLVFSAVVFGMAVQVSSPAVDYAKAKISAHLIMIEIKTFLDLSYSTGLMPT

LEGNVTFFGVFNYPTRPIPLVQLSLKLVKQTLALVSGSGCKGKSTVQLLERFYD

PLAGVLIDGKETKRLNVLWRAHLGIVSQEPLIFDCISAENIAGDNRVVSQBEIV

RAAKENIHAFTSLPNKYSTKVGDKGTOLSGGOKORIAALALVNPVKILLDEATS

ALDTESEKVVQEDLKAERSGRTCIIVIAHRLSTIYNADLIVVFNQGRVKEHGTHQQLA

QKGIYSMSVQAGTKRQ"

BASE COUNT 1372 a 892 c 1130 g 1252 t

ORIGIN

## Alignment Scores:

Pred. No.: 5,78e-222 Length: 4646  
Score: 3554.00 Matches: 682  
Percent Similarity: 75.63% Conservative: 246  
Best Local Similarity: 55.58% Mismatches: 263  
Query Match: 59.12% Indels: 36  
DB: 6 Gaps: 9

US-09-873-409-6 (1-1195) x AX391099 (1-4646)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 575 ATGTTGGGGAATTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATCGT 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCRAAATGACAGAAATTTAGAGATCTGATGCA 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43

Db 695 AACATCTAATAGAGATGATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 754

Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 755 ATGACAGGATGCTATTATACAGTGAATTTGGTCTGGGGTGGTGGTCTGCTTAC 814

Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81

Db 815 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAAGACAAATACACAAATTTAGAAAAACAG 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101

Db 875 TTTTTCATGCTATAATGCGACAGAGATAGGCTGTTTGTATGTCACGATGTTGGGAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120

Db 935 CTTAACACCCGACTTACAGATGATCTCTAAGATTAATGAAGGAATTTGGTGCACAAAT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140

Db 995 GGAAATGTTCTTCAGTCAATGCAACATTTTTCACATGGGTTTATAGTAGATTTACAGT 1054

Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160

Db 1055 GGTGTGAAGCTAACCCCTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTGACGTGCT 1114

Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180

Db 1115 GTCTGGGCAAGATATACTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

Db 1175 GGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAAGTGTGATTCATTTGGAGACAA 1234

Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220

Db 1235 AAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAG 1294

Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240

Db 1295 AAGACTTATACAGCAATATCTTCTAGTGTGCTTCTCTGCTGATCTATGATCATCTAT 1354

Qy 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260

Db 1355 GCTTGGCTCTCGATGGGACCACTTGGTCTCTCAGGGAA-----TATTCATTT 1408

Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280

Db 1409 GGAAGAGTACTCACTGATTTCTTTCTGATTAAATTTGGGGCTTTTGTGTCAGCAGGCA 1468

Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300

Db 1469 TCTCCAAGCATTTGAAGCATTTGCCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320

Db 1529 ATTGATAAATAGCAAGTATTGACAGCTATTGCAAGAGTGGGCAACAAACAGATAATATT 1588

Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340

Db 1589 AAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 1648

Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360

Db 1649 ATCTGAAGGGCTCAACCTGAAGGTGCAGAGTGGCAGACGGTGGCTGGTTGGAAAC 1708

Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380

Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGCTGATCAGAGGCTCTATGACCCACAGAG 1768

Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400

Db 1769 GGGATGTCAGTCTTGTGGACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGAA 1828

Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420

Db 1829 ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTTGCACACACAGTAGCTGAAACATT 1888

Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440

Db 1889 CGTATGGCCGTGAAATGTCCATGATGATGATGAGAAAGCTGTCAAGGAGGCAAT 1948

Qy 441 AlatyAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460

3029	Db	GCAATACGACGAGTGTGGTGAATGAAAAATGTTGCTGGACCAAGCACTGAAAGATAGAATAAGAAA	3088
807	Qy	GlulLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
3089	Db	GAACTAGAAGGTGCTGGGAAGATGCTGCTACTGAAGCAATAGAAAATCTCCGAACCGTGTCT	3148
827	Qy	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
3149	Db	TCCTTTGACTCAGAGACAGAAAGTTTGAAACATATGTATGTCTCAGAGTTCGCAGTACCATPAC	3208
847	Qy	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
3209	Db	AGAAACTCTTTGGAGAAAGCACACATCTTTGGATTACATTTCTCTCACCAGCAATG	3268
867	Qy	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleAlaGlyArg	886
3269	Db	ATGTATTTTTCCTATGCTGATGTTTCCGGTTTGGAGCCTACTTGTGGCACAATAAATC	3328
887	Qy	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
3329	Db	ATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTTCTTTGGTGCCATGCCGCTG	3388
907	Qy	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
3389	Db	GGGCAAGTCAGTTCATTTGCTCTGACTATGCCAAAGCCAAATATCAGACGCCACATC	3448
927	Qy	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro	946
3449	Db	ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGCCTAATGCCG	3508
947	Qy	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966
3509	Db	AACACATTGGAAGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCACCGACCG	3568
967	Qy	AspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrValAlaPhe	986
3569	Db	GACATCCAGTGCTTCAGGACCTGAGCCTGGAGGTGAAGAGGGCCAGACGCTGCTCTG	3628
987	Qy	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp	1006
3629	Db	GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCTCCAGCTCTCTGGAGCGTTCTACGAC	3688
1007	Qy	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	1026
3689	Db	CCCTTGGCAGGGAAGTCTGCTTGATGGCAAGAAATAAAGCCAGCTGAATGTTCCAGTGG	3748
1027	Qy	LeuArgSerGlnIleAlaIleValProGlnGluProValIlePheAsnCysSerIleAla	1046
3749	Db	CTCCGACGACACCTGGGCATCTGTGTCCTCCAGGAGCCCATCTCTGTTCAGTCAGCATTGTCT	3808
1047	Qy	GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluLysGluAla	1066
3809	Db	GAGAACATTGCCTATGGAGACACACGCCGGTGGTGTCTCAGAGAGAGATCGTGAGGCA	3866
1067	Qy	AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln	1086
3869	Db	GCAAAGGAGGCCAACATACATGCTTTCATCGACTCACTGCCCTAATAAATATAGCACTAAA	3928
1087	Qy	ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg	1106
3929	Db	GTAAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACACGATGTTGCCATAGCTCGT	3988
1107	Qy	AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn	1126
3989	Db	GCCCTTGTAGACAGCCTCATATTTTGGATTGATGAAGCCACGCTCAGCTCTGATACA	4048
1127	Qy	AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu	1146
4049	Db	GAAGTGAAGAGTGTGTCGAAGAGCCCTGGACAAAGCCAGAGAGAGGCCGCGACCTGCAT	4108
1147	Qy	ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn	1166
4109	Db	GTGATGTGCTACCGCTGTCCACCATCCAGATGTCAGATTAATAGTGTGTTCAGAAAT	4168

```
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4169 GGCAGAGTCAGGAGCATGCAGCATCAGCAGCTGCTGCACAGAAAGCACTATT 4228

QY 1187 LysLeuValAsnAlaGlnSer 1193
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4229 TCAATGTCAGTGCAGGCT 4249

RESULT 15
AF269224 4045 bp DNA linear SYN 04-JUL-2000
LOCUS Synthetic construct Canis familiaris his-tagged-multidrug
DEFINITION resistance glycoprotein gene, complete cds.
ACCESSION AF269224
VERSION AF269224.1 GI:8926216
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Analysis of dog MDR1 p-glycoprotein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Direct Submission
JOURNAL Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
Tournepieuille, BP3, Toulouse 31931, France
FEATURES
source
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="original mRNA isolated from Beagle dog colon"
101..3964
/codon_start=1
/transl_table=1
/product="his-tagged-multidrug resistance glycoprotein
MDR1"
/protein_id="AAF81747.1"
/db_xref="GI:8926217"
/translation="MFHHHHHDEGGKGSAEKNFWMKGSKKKNKKEKKPTVSTFA
MFYSNDRLLYMLVGTMAAIHGAALPLMLVFGNMTDSFANAGISRNKTFPVLIN
STNNTQHTLNEEMTVAAYYSIGIGAGVLVAAVYQVFWCLAAGRQILKIRKQFF
HLMROEIGFVDHVDGEINTRLTDDVSKINIGIGDKIMGFQSTATFTGIVGFTF
GKTLVILAIAPVLGSAIWAAILSSFTDKELLAYAKAGAAVEVLAIKRTVARG
GKLEKRNKNLEAKRGIGIKAITANISIGAFLLITASTALAFWYGTSLVLSSEY
SIGQVLTVFVSLGAFSGQASPSIEAFANARGAAVEIFPKIDNPKSIDSYSKSHK
PDNKGLEFKNVHFSPYRKEVKILKLNKVSQGTVALVNGSGKSTTVQLMQR
LYDPTGMVCIQDQIRTNVRLHREITGVVSOEPLVFATTAENIRYGRNVWDEI
EKAVKANAVDFTMKLPNKFDTLVGERGAQLSGGOKORIAIARALVRNPKILLDEAT
SALDTSEAVVQVALDKARKRTTIVIAHRLSTVRNADVIAGFDGDIIVEKGNHDELM
KEGIVFKLVMTQTRNEIELENATGESKESDALEMSPKDSGLSKRSTRSLHA
POQDRKGLGTEDLNEVPVPSFWRILKLNSTENVPYFVVGIFCAINGGLQAFSLIF
SRIGITFDEDEDPETRKSNMPSVLFLVGLIISFTFFLQGFQKAGEIILTKRLY
MVFPSLMRODVPFDDPKNTGALTFLANDAAOVKGAIGSLAVLTIONIANLGTGII
ISLIGWQTLILAIPLVIAIAGVEMKLSQALDKKELEGKGIATEALENPT
VLSLTREQFYMAYASLQVPRNSLRKAHIFGVSFSTIQAMWYFYACFPFGAYLV
ANEFMFPQDLVFSVAIVFGMAVQVSPFADYAKAVSAAHVIMILBSPLIDSYS
PHGLKENTGNTLVNEVFNVPTRDIPVLQGLSEVKKGTQALVSGSGCKSTV
QLLERYDPLAGSLVDGKEIKHLNWLRIHGVISOEPILFDCAIENIAYGNRSR
VVSHERIMOAQEKANTHFTETLPEKYNTRVGGKGTQLSGGOKORIAIARALVROPHI
LLIDEATSALDTESEKVVQEQALDKAREGRTCVIAHRLSTIQNADLIIVFQNGKVEH
GTHQLLAQGITFESMISVQAGKR"
BASE COUNT 1171 a 850 c 986 g 1038 t
ORIGIN

Alignment Scores:
Pred. No.: 6 54e-222 Length: 4045
Score: 352.00 Matches: 681
Percent Similarity: 74.74% Conservative: 239
Best Local Similarity: 55.32% Mismatches: 269
Query Match: 59.08% Indels: 42
```

Db	1283	AAATATTAAAGGGAATTTGGAAATTCAAAAATGTTCTACCTGATTACCCCTCTCGAAAAAGAA	1342
Qy	339	IleYsIleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
Db	1343	GTTAAGATCTTTAAGGGTCTCAACCTGAAGGTTACAGTGGGCAGACAGTGGCGCTGGT	1402
Qy	359	GlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
Db	1403	GGGAACAGTGGCTGCGGAAGACGACGCGGTGCAGCTGATGTCAGAGGCTCTATGACCCC	1462
Qy	379	AspAspGlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyr	398
Db	1463	ACAGATGGCATGCTGCTGATTGTATGGACAGGACATTAGGACCATAAATGTAAAGGCATCTT	1522
Qy	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
Db	1523	CGGGAATTAATTACTGGTGGTGAGTCAGGAGCCTGTGTGTTGCCACACCATAGCTGAA	1582
Qy	419	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu	438
Db	1583	AACATTGCTATGGCCGCGAAAAATGTCACCATGGATGAGATTGAGAAAAGCTGTTAAGGAA	1642
Qy	439	AlaAsnIaIaTyrAspPheIleMetGluPheProAenLysPheAsnThrLeuValGlyGlu	458
Db	1643	GCCAAATGCCATGATTTATCATGAACAACTACCTAAATTTGACACACTCGTGGTGGAGAG	1702
Qy	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	478
Db	1703	AGAGGGCCAGCTGAGTGGTGACAGAAACAGAGAAATGCCATTCCTCGGCGCCTGGT	1762
Qy	479	ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
Db	1763	CGCAACCCCAAGATTCTCTGCTGGATGAGGCAACGTCAGCTCTGCACACTGAAAGTGAA	1822
Qy	499	SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla	518
Db	1823	GCAGTGGTTGAGTGGCCCTGGATAAGGCGCAGAAAGGCGGACTACCATTTGTCATAGCT	1882
Qy	519	HisArgLeuSerThrIleArgSerAlaAepLeuIleValThrLeuLysAspGlyMetLeu	538
Db	1893	CATCGTTGTGTCAGTTCTGATGCGGATGCGGATGTCATTCGTGGTTTGATGGAGTCAAT	1942
Qy	539	AlaGluLysGlyAlaHisIaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal	558
Db	1943	GTGGAGAAGGAAATCATGATGAATCATGAAGAGAGAGGCACTTACTTCAAACCTGTC	2002
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
Db	2003	ACAATGTCAG--ACAGAGGAAATGAAATTGAAATTGAGAAATGCCACTGGTGAATCCAAA	2059
Qy	579	ArgLysThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle---	597
Db	2060	AGTGAAGTGTATGCTTGGAAATG---TCTCCAAAGATTACAGGTCAGTTTAATATAAA	2116
Qy	598	-----AspLysAlaGluLysUser	603
Db	2117	AGAAGATCAACTCGCAGGATATACATGCACACCAAGGCGCAAGACAGAAAGCTTGGTACA	2176
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAen	623
Db	2177	AAAGAGGACTTTGAATGAGAATGTACCTCCAGTTTCTCTCGAGGATTCGAAAGCTGAAC	2236
Qy	624	LysProGluIleTyrProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrVal	643
Db	2237	TCAACTGAATGGCCCTATTTTGGTGGTGTATATTTTGGCTGATTATAAACGGAGGCCCTG	2296
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	662
Db	2297	CAACGACATTTTCATTAATATTTTCAGGATTTATAGGATCTTTTACC CGAGATGAGAT	2356
Qy	663	LysThrLeuTyrHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682

Db	2357	CCTGAAACAAAACGACAGAAATAGTAACATGTTTTCTGTGATTATGTTCTTAGTCCTTGGAATT	2411
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2417	ATTTCTTTTATTACATTTTTTCTCCAGGGCTTCACATTTGGCAAAAGCTGGGAGATCCTC	2476
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2477	ACTAAGCGCTTCGATACATATGTTTTCAGATCCATGCTGAGACAGGATGTCAAGCTGGTTT	2536
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2537	GATGACCTTAAACACACCACTGGAGCATTTGACACAGAGCTTGCCAAATGATCGGCTCAA	2596
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2597	GTTAAAGGGGCTATAGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAAATCTTGGG	2656
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheIleuIleLeuSerIle	782
Db	2657	ACAGGCATTTATTATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT	2716
Qy	783	AlaProValIleAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2717	GTACCCATCATTTGCAATACAGAGAGTTGTGTGAATAATGTTGTCTGGCAAGCACTG	2776
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
Db	2777	AAAGATAAGAAAGAGCTAAGAGAGCTGGAGATTTGCTACAGAAGCCATCGAAAACTTC	2836
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2837	CGAACTGTGTCTTCTTGACTCGGAGCAGAAAGTTTGAATACATGATGTATGCACAGAGTTT	2896
Qy	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	862
Db	2897	CAAGTACCATACAGAAACTCTTTGAGAAAGCACACATCTTCGGGTCTCATTTCTATC	2956
Qy	863	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	882
Db	2957	ACCCAGGCAATGATGATTTTTCCTATGCTGGCTGTTTCGGGTTTGGTGCTACTTGGTG	3016
Qy	883	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly	902
Db	3017	GCAAAATGAGTTTCATGAACCTTTCAGGATGTTCTTTTGGTATTTCTCAGCTATGTCTTGGT	3076
Qy	903	AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	922
Db	3077	GCCATGGCATGGGCGAGTCAGTTCATTTGCTCTCGACTATGCCAAGCAAGATATCA	3136
Qy	923	AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu	942
Db	3137	GCAGCCACGTCATCATGATCATTTGAAAAAGGCCCTCTGATTGCACAGCTACAGCCCTCAC	3196
Qy	943	GlyLysGlyProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	962
Db	3197	GGCCTCAAGCCAAATACGTGTGAAGGAATAATGTGACATTTAATGAGGTCGTGTCAACTAT	3256
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLysLeuSerLeuSerIleGluArgGlyLys	982
Db	3257	CCCACTCGACACATCCCCGTCGTCAGAGGGCTGAGCCTCGAGGTGAAGAGGGCCAG	3316
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln	1002
Db	3317	ACGCTGGCCCTCGTAGTAGCAGTGGCTGTGGGAAGACACAGTTGTTACGCTCCTAGAG	3376
Qy	1003	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3377	CGCTTCTATGACCCCTTGGCTGGTTTCAGTGTCTAATTGATGGCAAGAGATAAAGCACCTG	3436
Qy	1023	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3437	AATGTCCAGTGGCTCCGAGCACACCTCGGCATCTGCTCAGAGGCCATCTCTGTTTGG	3496

```

QY 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGlu 1062
|||
|||
|||
Db 3497 TGCAGCATTCGCCGAGACATTCCTATGGAGACACAGCCGGGTCTGTATCACATGAAGAG 3556
|||
|||
|||
QY 1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
|||
|||
|||
Db 3557 ATTATGCGAGCAGCAGGAGGCCAACATACACCCTTCATCGAGACACTCCCTGAGAAA 3616
|||
|||
|||
QY 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
|||
|||
|||
Db 3617 TACAACACCCAGAGTAGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT 3676
|||
|||
|||
QY 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 1122
|||
|||
|||
Db 3677 GCCATAGCTCGCGCTCTTGTAGACAGCCTCATATTTTGGTGGATGAAGCTACATCA 3736
|||
|||
|||
QY 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
|||
|||
|||
Db 3737 GCTCTGGATACAGAAAGTGAAGGTTGTCCAGAGAGCCCTGGACAAAGCCAGAGAGGC 3796
|||
|||
|||
QY 1143 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
|||
|||
|||
Db 3797 CGCACCTGCATTGTATGCGCCACCGCTTGTCCACCATCCAGATGCAGATTTAATAGTG 3856
|||
|||
|||
QY 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlnLeuLeuArgAsnArg 1182
|||
|||
|||
Db 3857 GTGTTTCAGATGGCAAGTCAGAGAGCATGGCAGATCAACAGCTGCTGGCTCAGAAA 3916
|||
|||
|||
QY 1183 AspileTyrPheLysLeuValAsnAlaGlnSer 1193
|||
|||
|||
Db 3917 GGCATCTATTTTCCATGATCAGTGTCAGGCT 3949
|||
|||
|||

```

Search completed: March 31, 2003, 02:22:27  
Job time : 16406.1 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 1167.91 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012

Sequence: 1 MILGILASLVNGACILMPL.....QELLNRDIYFKLVNAQSVQ 1195

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q/cn2\_1/USPTO.spool/US09873409/runat\_27032003\_115416\_19216/app\_query.fasta\_1.7544  
-DB=N.GeneSeq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 4247 @runat\_27032003\_115416\_19216 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: N.GeneSeq\_101002.\*
- 2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT.\*
- 3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT.\*
- 4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT.\*
- 5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT.\*
- 6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT.\*
- 7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT.\*
- 8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT.\*
- 9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT.\*
- 10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT.\*
- 11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT.\*
- 12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT.\*
- 13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT.\*
- 14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT.\*
- 15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT.\*
- 16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT.\*
- 17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT.\*
- 18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT.\*
- 19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT.\*
- 20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT.\*
- 21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT.\*
- 22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT.\*
- 23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT.\*
- 24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT.\*
- 25: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	5760.5	95.8	3699	24	ABK83223	Human transporter
2	3557	59.2	4279	22	AAO03506	Dog P-glycoprotein
3	3556.5	59.2	4195	22	AAF86128	Cynomolgus monke
4	3554	59.1	3860	21	AAZ49332	Human wild-type mu
5	3554	59.1	3860	24	ABA94365	Human BCRP DNA rel
6	3554	59.1	4279	22	AAO03504	Dog P-glycoprotein
7	3554	59.1	4279	22	AAO03505	Dog P-glycoprotein
8	3549	59.0	4279	22	AAO03488	Dog P-glycoprotein
9	3549	59.0	4317	22	AAO03489	Dog P-glycoprotein
10	3548	59.0	4186	22	AAF86127	Cynomolgus monke
11	3545	59.0	3860	21	AAZ49333	Human G185V mutant
12	3545	59.0	3860	24	ABA94366	Human BCRP DNA rel
13	3545	59.0	4349	22	AAH57442	Human intestine ce
14	3545	59.0	4646	21	AAZ94738	Human ATP binding
15	3545	59.0	4646	24	AAZ38994	Human mdrl gene.
16	3545	59.0	4646	24	ABL68592	Kidney cancer rela
17	3545	59.0	4646	24	ABL68880	Kidney cancer rela
18	3545	59.0	5505	17	AAH13394	Hybrid vector pSF-
19	3545	59.0	8630	21	AAZ24041	Retroviral M4 mdr-
20	3545	59.0	8630	21	AAZ24042	Hybrid vector pSF-
21	3543.5	58.9	4264	19	AAV66533	Mutated human P-gl
22	3543.5	58.9	4264	19	AAV66534	Mutated human P-gl
23	3543	58.9	4669	8	AAH70752	Sequence of human
24	3540.5	58.9	4425	21	AAZ52048	Rat multidrug resi
25	3539	58.9	4378	11	AAQ04522	Multidrug Resistan
26	3538.5	58.9	4369	21	AAZ52047	Rat multidrug resi
27	3537.5	58.8	3840	24	ABL91687	Human polynucleoti
28	3537.5	58.8	3988	21	AAZ88973	Human MDR-1 DNA.
29	3536	58.8	4669	14	AAQ52726	Sequence of human
30	3534	58.8	4646	15	AAQ72872	Human multidrug re
31	3527	58.7	4669	21	AAV32645	Human P glycoprote
32	3526.5	58.7	4189	21	AAZ49334	cDNA encoding huma
33	3526.5	58.7	4189	21	AAZ49331	Murine multidrug r
34	3526.5	58.7	4189	24	ABA94367	Mouse BCRP DNA rel
35	3526.5	58.7	4313	14	AAQ38950	Mouse multidrug re
36	3522.5	58.6	4788	21	AAZ49335	Murine multidrug r
37	3522.5	58.6	4788	24	ABA94368	Mouse BCRP DNA rel
38	3517.5	58.5	4233	21	AAZ90198	Rat mdrlb2 (multis
39	3517.5	58.5	4233	22	AAF27498	Rat mdrlb2 multidr
40	3504.5	58.3	3924	21	AAZ94742	Human ATP binding
41	3504.5	58.3	3924	21	AAZ88974	Human MDR-3 DNA.
42	3504.5	58.3	3924	24	ABN95801	Gene #2299 used to
43	3476.5	57.8	4254	24	ABK63517	Rat sequence diffe
44	3428.5	57.0	3912	24	ABK63653	Rat sequence diffe
45	2868	47.7	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1

ABK83223

ID ABK83223 standard; cDNA; 3699 BP.

XX ABK83223;

XX 27-AUG-2002 (first entry)

XX Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.

XX Human; ss; gene; transporter and ion channel; TRICH; transport disorder;  
KW neurological disorder; muscle disorder; immunological disorder; cancer;  
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;  
KW cell proliferative disorder; cervical cancer; breast cancer;  
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;  
KW myotonic dystrophy; catatonia; endocrine disorder; diabetes;  
KW Grave's disease; gastrointestinal disorder; Crohn's disease;  
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;  
 KW protozoal infection; helminthic infection; cardiovascular disorder;  
 KW atherosclerosis; hepatic disease.

XX Homo sapiens.

XX W0200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-NOV-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-250790P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

XX Rankumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for  
 PT diagnosing, preventing, and treating disorders associated with an  
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular  
 PT or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
 CC a recombinant polynucleotide comprising a promoter sequence operably  
 CC linked to the TRICH polynucleotide, a cell transformed with the  
 CC recombinant polynucleotide, a transgenic organism comprising the  
 CC recombinant polynucleotide, an isolated antibody that binds specifically  
 CC to TRICH, and screening for compounds which bind to TRICH, modulate  
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.  
 CC The polypeptides are useful for diagnosing, treating, and  
 CC preventing transport, neurological, muscle, immunological disorders  
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or  
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
 CC catatonion), endocrine disorders (e.g. diabetes, Grave's disease),  
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,  
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
 CC other diseases and disorders detailed in the specification. They can also  
 CC be used in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of transporters and  
 CC ion channels. TRICH or its fragments may also be used in screening for  
 CC compounds that specifically bind to and modulate the activity of TRICH.  
 CC The polynucleotides can be used to create knock-in humanised animals or  
 CC transgenic animals to model human disease. The present sequence  
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

XX Alignment Scores:

Pred. No.: 0 Length: 3699

Score: 5760.50 Matches: 1156

Percent Similarity: 95.87% Conservative: 4

Best Local Similarity: 95.54% Mismatches: 9  
 Query Match: 95.82% Indels: 41  
 DB: 24 Gaps: 3

US-09-873-409-6 (1-1195) x ABK83223 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 145 ATGATCCTGGGTATACCTGACATCAGTGTTCATGGAGCCTGCTTCTTAAAGCCACTG 204  
 QY 21 ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
 DB 205 TGTATAGGAGAAATGATGATTAACCTTATTAGTGGATGTCTAGTCCACACACAACA 264  
 QY 41 SerPhePheArg-----LeuThr 46  
 DB 265 AATTATCAGAACTGTACTCAGTCTCAAGAGAACTGAATGAAGATATGACTCTGTGTGACC 324  
 QY 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
 DB 325 CTGTATTATGTTGGAATAGGTGTGCTGCTTGTATTTTGGTTACATACAGATTCTCTTG 384  
 QY 67 TrpIleIleThrAlaAlaArgGlnThrIleArgIleAArgGlyGlnPhePheHisSerVal 86  
 DB 385 TGGATTATAACTGCAGCAGCAGACCAAGAGATTCGAAAAACAGTTTTTTCATTTCAGTT 444  
 QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
 DB 445 TTGGCAGCAGACATCGCTGGTGTGATAGCTGTGACATCGGTGAACCTTAACACTCGCATG 504  
 QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
 DB 505 ACAGATGACATTGACAAATCAGTGTATGATGATGATGATGATGATGATGATGATGATG 564  
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValGlySerGlyTrpLysLeuThr 145  
 DB 565 AACATGCTACTTTTTTCGATTGGCTGGCAGTTGGTTTGGTGAAGGCTGGAAACTCACC 624  
 QY 146 LeuValThrIleuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165  
 DB 625 CTAGTAGTACTCTATCCACAGTCTCTCTTATTAATGGCTTCAGCGCAGCATGTTCTAGATG 684  
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
 DB 685 GTCACTCTATTGACCACTAAGGNAATTAAGTGCCTATTCCAAAGCTGGGCTGGGCGAA 744  
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205  
 DB 745 GAAGTCTTGTCTCAATCCGAAACAGTCATAGCTTTAGGGCCCGAGGAGAAAGAACTTCAA 804  
 QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225  
 DB 805 AGGTATACACAGAAATCTCAAGAGATGCAAGAGATTTTGGCATATAAAGGACATATAGCTTCA 864  
 QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyr 245  
 DB 865 AAGTGTCTCTTGGTCTGTGTACTCTTTATGAATGGAACCTATGACTTGTCTTTTGG 924  
 QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265  
 DB 925 TATGGAACCTCTTGTATCTTAATGGAGAACCTGGATATACCATCGGAGCTGTCTTGTCT 984  
 QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
 DB 985 GTTTTCTTTAGTGTAAATCCATAGTAGTTATTCATTTGGAGCAGCAGTCCCTCACTTTGAA 1044  
 QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305  
 DB 1045 ACCTTGCATATGCCCGAGGAGCTGCTTTCATATTTTCCAGGTATTTATATAGAACC 1104  
 QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325  
 DB 1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGAACTGTGGAA 1164



Db 3247 GCTGAGAACATCGCTTATGCTGCAACAGCGGTGTGTGTCATTAGATGAGATCAAGAA 3306  
Qy 1066 AlaAlaAsnAlaAlaAenIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085  
Db 3307 GCCGCAATGCAGCAATATCCATTCTTTATTAAGGGTCTCCCTGAGAAATCAACACA 3366  
Qy 1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105  
Db 3367 CAATGTGGACTGAAGAGGACACAGCTTCTGCGCGCCAGAAACAAAGACTAGCTATTGCA 3426  
Qy 1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125  
Db 3427 AGGCGCTCTCTCCAAAACCCCAAAATTTATTGTTGATGAGGCCACTTCAGCCCTCGAT 3486  
Qy 1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 1145  
Db 3487 AATGACAGTGAGAAGGTGTTTCAGCATGCCCTTGATAAGCCAGGACGGAGGACATGC 3546  
Qy 1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165  
Db 3547 CTAGTGTCTACTACAGGCTCTCTGCAATTCAGAACGCAGATTGTAGTGTCTGCGAC 3606  
Qy 1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185  
Db 3607 AATGGAAGATAAAGGAACAAGCACTCATCAAGAGCTCTGAGAAATCGAGACATAT 3666  
Qy 1186 PheLysLeuValAsnAlaGlnSerValGln 1195  
Db 3667 TTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696

## RESULT 2

AD03506  
ID AAD03506 standard; cDNA; 4279 BP.

XX  
AC  
XX  
XX  
DT

13-JUN-2001 (first entry)

DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.

XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
KW MDR1, drug bioavailability; transgenic animal; genetic model; ss.

XX Canis familiaris.

XX Key Location/Qualifiers  
FH 17..3862  
CDS

FT /\*tag= a  
FT /product= "Dog P-glycoprotein (PGP) allelic variant  
FT (Genotype D) protein"

FT allele

FT replace (91, T)

FT allele

FT replace (607, C)

FT allele

FT replace (1001, T)

FT allele

FT replace (3458, A)

FT /\*tag= c

WO200123540-A2.

XX

PD 05-APR-2001.

XX

PF 28-SEP-2000; 2000WO-US26767.

XX

PR 28-SEP-1999; 99US-0156510.

XX (GENT-) GENTEST CORP.

XX

PI Stocker PJ, Steinel-crespi DT, Crespi CL, Reif TC, Patten CJ;

XX WPI; 2001-235373/24.

DR

DR P-PSDB; AAE00310.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT PGP inhibitors -

XX Claim 9; Page 102-107; 11pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity, by  
CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump  
CC exporting small molecules across the cell membrane. This enzyme  
CC is a member of the ABC transporter family.

XX Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 4279  
Score: 3557.00 Matches: 683  
Percent Similarity: 74.74% Conservative: 237  
Best Local Similarity: 55.48% Mismatches: 269  
Query Match: 59.17% Indels: 42  
DB: 22 Gaps: 8

US-09-873-409-6 (1-1195) x AAD03506 (1-4279)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20  
Db 167 ATGTGTGGGACAAATGCTGCCATCATCCATGAGCTGCACCTCCTCTCATGATGCTG 226  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 227 GTTTTGGAAACATGACAGATAGCTTTCGAATGCGAGGAATTCAGAAACAAACTTTT 286  
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 287 CCAGTTTATTAATTAATGAAGTATTACGAACAATACACACATTTTCATCAACCATCTGGAG 346  
Qy 45 -----LeuThrLeuTyrTyrValGlyValAlaAlaLeuIlePhe 59  
Db 347 GAGGAAATGACCACGTATGCCCTATTATTACAGTGGATCGGTGCTGGCGTGTGGGCT 406  
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleThrAlaAlaAArgGlnThrLysArgIleArg 79  
Db 407 GCTTACATCCAGTTTCATTCTGTCCTGGCAGCAGGAGACAGATATCAAAATTAGA 466  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
Db 467 AAACAATTTTTTCATGCTATCATCGACAGGAGATGGCTGGTTGACGTGATGACGTT 526  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 527 GGGGAGCTTAACACCGCGCTCACAGCATGTCTCCAAATCAATGAAGAAATGCGGAC 586  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 587 AAAATTGGAATGTTCTTCAATCAATAGCAACATTTTCCCGGTTTATAGTGGGTTT 646  
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 647 ACACGTGGTTGGAAGCTAACCCCTTGTGATTGTTGGCCATCAGCCCTGTTCTTCTGACTTTCA 706

Qy	159	AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer	178
Db	707	CGCGCCATCTGGCCAAAGATACTATCTTCATTACTGATAAAGAACTCTTGGCCTATGCA	766
Qy	179	LysAlaGlyAlaValaGluGluValLeuSerSerIleArgThrValIleAlaPheArg	198
Db	767	AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGACTGTGATGGCTTTGGG	826
Qy	199	AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	218
Db	827	GGACAAAGAAAGAACTCTGAAAGGTACAACAAATTTAGAAGAAGCTAAAGGAATTGGG	886
Qy	219	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	238
Db	887	ATAAAGAAAGCTATCAGCGCCACATTTCTATTGTGGCGCTTCTTATTGATCATATGCA	946
Qy	239	ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	258
Db	947	TCATATGCTCTGGCTTCTGGTATGGAGCTCTTGGTCTCTCTCAGTGAA-----TAT	1000
Qy	259	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	278
Db	1001	ACTATTGGACAAGTACTCACTGCTCTTTCTGATTAAATTGGGGCTTTTAGTATTGGA	1060
Qy	279	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	298
Db	1061	CAGGCATCCCACAGCATTTGAAGCATTTGCMAACGCAAGAGGAGCAGCTTATGAAATCTTC	1120
Qy	299	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	318
Db	1121	AGATAATTGCAATTAACACAGCATTTGACACTATTGGAAGGTGGACATTAACACGAT	1180
Qy	319	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	338
Db	1181	AAATATTAAAGGAAATTTGGAAATTCAAAAATGTTCACTTCAGTTACCTCTTCGAAAAAGAA	1240
Qy	339	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	358
Db	1241	GTTAAGATCTTTAAAGGCTCTCAACTGAAGGTTCAGAGTGGGCGACAGATGGCGCTGGTT	1300
Qy	359	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro	378
Db	1301	GGGACACAGTGGCTCGGGAAGAGCACGCCGTGCAGCTGATGCAGAGGCTCTATGACCCC	1360
Qy	379	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	398
Db	1361	ACAGATGGCATGCTGTGATTGATGCACAGGACATTAGGACCAATAATGTGAAGGCATCTT	1420
Qy	399	ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	418
Db	1421	CGGGAAATTACTCGTGTGGTAGTCAGGAGCTGTGTTGTTGGCCACACAGATAGCTGAA	1480
Qy	419	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu	438
Db	1481	AACATTGCCTATGGCCGCGAAAATGTCAACATGGATGAGATTGAGAAAGCTGTTAAGGAA	1540
Qy	439	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	458
Db	1541	GCCAAATGCCCTATGATTTTATCATGAAACTACCTTAATAAATTTGACACTCTGTTGGAGAG	1600
Qy	459	LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	478
Db	1601	AGAGGGCCACGCTGAGTGTGTGCAGAAACAGAGAAATCGCCATTGCTCGGGGCCCTGGTT	1660
Qy	479	ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys	498
Db	1661	CGCAACCCCAAGATTCTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA	1720
Qy	499	SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValAla	518
Db	1721	GCAGTGGTTCAGTGGCCCTGGATAGGCCAGAAAAGGCCGACTACCATTTGTGATAGCT	1780

Qy	519	HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu	538
Db	1781	CATCGTTGTGTACAGTTCGTAAATCCGGATGTCATTCGTGCTGTTGTGATGATGGAGTCATT	1840
Qy	539	AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysAspGlyLeuTyrTyrSerLeuVal	558
Db	1841	GTGGAGAAAGGAATCATGATGAACATCATGAAGAGAGAGGCACTTTACTTCCAAACTTGTCT	1900
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
Db	1901	ACAATGCAG--ACAAGAGAAATGAATTTAGTTAGATAATGCCACTGGTGAATCCAAAC	1957
Qy	579	ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle---	597
Db	1958	AGTGAAGAGTCAATCCCTGGAAATG---TCTCCAAAAGAGTTACAGGTCACAGTTTAAATAAA	2014
Qy	598	-----AspLysAlaGluGluSer	603
Db	2015	AGAATCAACTCGCAGGAGTATACATGCACCAAGGCCAAGACAGAACTTGGTACA	2074
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
Db	2075	AAAGAGGACTTGAATCAGAAATGTAACCTCCAGTTCCTCTCGAGGATTTCTCAAGCTGAAC	2134
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2135	TCAACTGAATGGGCTTATTTTGTGGTGTGATATATTTGTGCTATTTATATAACCGAGGCGCT	2194
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsp	662
Db	2195	CAACGACGATTTTCAATAATATTTTCAAGGATATATAGGATCTTTACCCGAGATGAGGAT	2254
Qy	663	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2255	CCTGAACAAACACACAGAATAGTAACATGTTTCTGTATGTTCTTAGTCTCTTGGAAAT	2314
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2315	ATTTCTTTTATTAACATTTTCTCCAGGCTTCACATTTTGGCAAGCTGGGGAGATCCTC	2374
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2375	ACTAAGCGGCTTCGATACATGTTTTCAGATCCATGCTGACAGAGATGTCAAGTGGT	2434
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2435	GATGACCTTAANAACACCACCTGGAGCATTTGACACACGAGTTCGAATGATCGCGCTCA	2494
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2495	GTTAAAGGGCTATAGTTCGAGGCTTGCTGTCTCATTCACCAATATAGCAAACTCTTGGG	2554
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle	782
Db	2555	ACAGGCATTTATATCCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAAT	2614
Qy	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2615	GTACCCATCATTCGATAGCAGAGTGTGTAATGAATGAATGTTCTCTGGACAGCACTG	2674
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
Db	2675	AAAGATAAGAAAGAGCTAGAAGAGAGCTGGGAAGATTGCTACAGAGCCATCGAAACCTC	2734
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2735	CGAATCTGTTGTTCTTTGACTCGGGAGCAGAAAGTTGAATACATGATGATGCACAGAGTTTG	2794
Qy	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe	862
Db	2795	CAAGTACATACAGAAACTCTTTGAGGAAGACACACATCTCTGGGGTCTCATTTCTATC	2854
Qy	863	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyValatyrLeuIle	882





[illegible]

1384	Db	GTTCGAAACACAGCGCTGTGGGAAGACCAACGCTGCAGCTGATCCAGAGGCTTTATGAC	140
378	Qy	ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValargHis	397
1444	Db	CCACACAGGGCATGGTTCAGTGTGTGATGCACAGGATATTAGGACCATAAACGTAAAGTTT	1503
398	Qy	TyrzhrAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer	417
1504	Db	CTACGGGAAATCATCGTGTGGTGAAGTCTAGGAACTGTTATGTTTGGCCACACGATAGCT	1563
418	Qy	AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg	437
1564	Db	GAATAACATTCCGTATGGTCTGAGATGTCCACATGGATGAGATTGAGAAGAGCTGTCAAG	1623
438	Qy	GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly	457
1624	Db	GAAGCCAAATGCCATTATGCTTTATCATGAACATGCCCTCAGAAATTTGCACACCTGGTTGGA	1683
458	Qy	GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu	477
1684	Db	GAGAGAGGGCCACAGCTGAGTGGTGGCGAAGACGAGAGGATGCCCATTTGCACGTGCCCTG	1743
478	Qy	ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer	497
1744	Db	GTTTCGAAACCCCAAGATCTCTCTGCTGGACGAGGCCAGCTGACGCTTCGACACGAAGAAT	1803
498	Qy	LysSerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal	517
1804	Db	GAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAGGTCGACACCATTTGTGTA	1863
518	Qy	AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet	537
1864	Db	GCTCATCGTTTGTCTACGGTTCGTAATCGGACGTCATCGCTGGTTTCGATGATGGAGTC	1923
538	Qy	LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu	557
1924	Db	ATTCTGGAAGAAGAAATCATGATGAGCTCATGAAGAAAGAGCATTTTACTTTCAAACTT	1983
558	Qy	ValMetSerGlnAspIle-----LysLysAlaAspGlu-----	568
1984	Db	GTCACAATGCAGACAGCAGGAATGAAATTGAATAGAAAATGCAGCTGATGAATCCAA	2043
569	Qy	---GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu-----	584
2044	Db	AGTGAAATTGATACCTTGGAAATGTCTTCATCATGATTTCAGGATCCAGTCTAATAGAAAA	2103
585	Qy	-----ProLeuHisSerValLysSerIleLysSerAppPheIleAspLysAlaGlu	602
2104	Db	AGATCCACTCGTAGAGTGTCGCTGGATCACAAGGCCAA-----GACAGAAAGCTTAGT	2157
603	Qy	SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeu	622
2158	Db	ACCAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCTCTTTGGAGGATTATGAGCTA	2217
623	Qy	AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr	642
2218	Db	AATTAACTCAGTGGGCTTATTTTGTGTGTGTATTTTGTGCCATTATAAATGGAGGT	2277
643	Qy	ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsn	661
2278	Db	CTGCAACCCACGATTTCCGCTAATATTTTCCAAAGATATATAGGGATTTTACAGAAGATGAT	2337
662	Qy	AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly	681
2338	Db	GATCCCAAAACAAACGACGAATAGTAACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2397
682	Qy	ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyIle	701
2398	Db	ATTGTTCTTTTATTACATTTTTCCTTCAGGCTTCACATTTTGGCAAAAGCTGGAGATC	2457
702	Qy	LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp	721
2458	Db	CTCACCAAGGGGCTCGATACATGATGTTTTTCCGATCCATGCTCAGACGAGATGTGATCGG	2517



QY 722 PheAspGluLysGluAenSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 741  
 DB 2518 TTTGATGACCCCTAAACACACCATCGAGCATTCAGTACAGGCTCGCCATGATGCTGCT 2577  
 QY 742 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAenAlaThrAenMet 761  
 DB 2578 CAAGTTAAAGGGCTATAGGTTCCAGGCTTGTATATATACCCAGAAATATAGCAATCTT 2637  
 QY 762 GlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSer 781  
 DB 2638 GGGACAGGAATAATATATCTTAATCTATGTTGGCACTGACACTGTTACTCTTATAGCA 2697  
 QY 782 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 801  
 DB 2698 ATTGTACCCATCATTCGAATAGCAGGAGTGTGTAATGAATGAAATGTTGTCTGCACAAGCA 2757  
 QY 802 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAen 821  
 DB 2758 CTGAAGATTAAGAAAGAACTAGAAGGTGCTGGGAAGATCGTACTCTGAAGCAATAGAAAAC 2817  
 QY 822 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 841  
 DB 2818 TTCGGAACCTGTTCTTTGACTCAGGAGCAGCAAGTTCGAACATATGTATGATCAGAGT 2877  
 QY 842 LeuGlnThrGlnHisArgAenThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 861  
 DB 2878 TTCAGGTACCATACAGAACTCTTTGAGGAAGACACACATCTTTGGAATCAGTTTTC 2937  
 QY 862 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 881  
 DB 2938 TTCACGAGGCAATGATGATTTTTCATATGCTGATGTTTCCGTTTGGAGCTACTTG 2997  
 QY 882 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 901  
 DB 2998 GTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTGTATATTTTCAGCTGTGCTTT 3057  
 QY 902 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 921  
 DB 3058 GTTGCCATGCGCCGTGGGCAAGTCAGTTCTATGCTGCTGACTATGCAAGCAAGCAAGTA 3117  
 QY 922 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAenIleAspSerArgSerGln 941  
 DB 3118 TCAGCAGCCCATCATCATCATTCATTTGAAAAAACCCCTTTGATTCAGACTACAGCACA 3177  
 QY 942 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 961  
 DB 3178 GAAGGCCCTAAAGCCGAACATTTGGAAGGAATGTCTCATTTAATGAAGTTGTATTCAAC 3237  
 QY 962 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 981  
 DB 3238 TATCCCAACCCGACTGGACATCCCGAGTCTTCAGGGGCTGAGCCTGGAGTGAAGAGGGC 3297  
 QY 982 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 1001  
 DB 3298 CAGAGCTGCGCCCTGTGGGCGAGCAGTGGCTGTGGGAAGAGCAGCGGTGGTCCAGCTCCTG 3357  
 QY 1002 GlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGlu 1021  
 DB 3358 GAGCGGTCTATGACCCCTTTGGCGGGGAAGTGTCTTGACGGCAAGAAATAAAGCAA 3417  
 QY 1022 LeuAenValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 1041  
 DB 3418 CTGATGTTTCAGTGGCTCCGAGCACACTGGGCATCTGTGTCAGGAGCCATCTCTGTTT 3477  
 QY 1042 AsnCysSerIleAlaGluAenIleAlaTyrGlyAspAenSerArgValValProLeuAsp 1061  
 DB 3478 GACTGCAGCATTTAGTGAACATTTGCTATGGAGACAACAGCCGGGTGGTGTCCACAGAA 3537  
 QY 1062 GluIleLysGluAlaAlaAenAlaAenIleHisSerPheIleGluGlyLeuProGlu 1081  
 DB 3538 GAGATCGTGGGCGAGCAAGAGGGCCCAATATACACGCTTTCATCGAGTCACTGCCTAAT 3597

QY 1082 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 1101  
 DB 3598 AAATATAGCACCCAGAGTAGGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAACGC 3657  
 QY 1102 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 1121  
 DB 3658 ATTGGCATAGCTCGTGCCCTTTAGACACAGCTCATATTTTCTTTTGGATGAGCCACA 3717  
 QY 1122 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1141  
 DB 3718 TCAGCTCTGGATACAGAAAGTGAAGGTTGTCGAAGAAGCCCTGGACAAAGCCAGAA 3777  
 QY 1142 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1161  
 DB 3778 GGCGGTACCTGCTGATGTGCTCACCCTGTCCCATCCAGATCCAGACTTAATA 3837  
 QY 1162 ValValLeuHisAsnGlyLysIleLysGlnGlnGlyThrHisGlnGluLeuLeuArgAsn 1181  
 DB 3838 GTGGTGTTCGAAGTGGCAGATCAAGGAGCAGCGCACATCAGCAGCTGCTGGCACAG 3897  
 QY 1182 ArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193  
 DB 3898 AAAGCATCTATTTTCAATGCTGCTCAGTGTCCAGGCT 3933  
 AC AAZ49332;  
 XX 14-MAR-2000 (first entry)  
 XX Human wild-type multidrug resistance-1 (MDR-1) cDNA.  
 DE Multidrug resistance; MDR-1; P-glycoprotein;  
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassaemia;  
 KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
 KW cytokine; wild-type; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FH CDS 1..3843  
 FT /\*tag= a  
 FT /product= "Human wild-type MDR-1 protein"  
 FT /replace (553..555, GTT)  
 FT /\*tag= b  
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given in AAZ49333"  
 PN WO9961589-A2.  
 XX 02-DEC-1999.  
 PD 27-MAY-1999; 99WO-US11825.  
 PF 28-MAY-1998; 98US-0086988.  
 PR (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA Sorrentino B, Bunting K;  
 XX WPI; 2000-072615/06.  
 DR P-PSDB; AAY58186.  
 XX Ex vivo expansion of hematopoietic stem cells transduced with a  
 PT sequence encoding human multidrug resistance-1, used for bone marrow  
 PT transplantation -  
 XX Claim 10; Page 68-70; 113pp; English.  
 PS

CC This sequence represents cDNA encoding human wild-type  
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane  
 CC efflux pump, responsible for the export of drugs from cells,  
 CC particularly cancer cells. Wild-type MDR-1 shows increased  
 CC resistance to etoposide and decreased resistance to vinca  
 CC alkaloids compared with a mutant form (AY58187) where the Gly at  
 CC position 185 is replaced by Val. The invention relates to transducing  
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified haematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC haematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,  
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in haematopoietic stem cells.  
 CC Haematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC haematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.

XX Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.:	0	Length:	3860
Score:	3554.00	Matches:	682
Percent Similarity:	75.63%	Conservative:	246
Best Local Similarity:	55.58%	Mismatches:	263
Query Match:	59.12%	Indels:	36
DB:	21	Gaps:	9

US-09-873-409-6 (1-1195) x AA249332 (1-3860)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	151	ATGGTGGTGGAACTTTGGCTGCCATCCATCGGGCTGGACTTCCTCTCATGATGCTG	210
Qy	21	ValLeuGlyGluMetSerAsp-----AsnLeuIleSer	31
Db	211	GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGAAATTTAGAAAGATCTGATGTC	270
Qy	32	GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-----	43
Db	271	AACATCACTAATAGAGATGATATCAATGATACAGGGTCTTTCATGATCTGGAGGAGAC	330
Qy	44	-----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr	61
Db	331	ATGACCAGATATCCCTATTATTACAGTGAATTTGGTGTGGGTGCTGCTTAC	390
Qy	62	IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleArgIleGln	81
Db	391	ATTTCAGGTTTCATTTTGGTGGCTGGAGCTGGAGAACAAATACACAAATTTAGAAACAG	450
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101
Db	451	TTTTTTCATGCTAATATCCAGACAGAGATAGGCTGGTTGATGTCACGATGTTGGGAG	510
Qy	102	LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspIle	120
Db	511	CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAATGAAGGTATTGGTGACAAAT	570
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	571	GGAATGTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT	630
Qy	141	GlyTrpIleLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	631	GGTTGAAGCTAACCTTGTGATTTTGGCCATCAGTCCCTGCTTCTTGACGTGTCAGCTGCT	690
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180

Db	691	GTCTGGGCAAGATATACTATCTTCACTTACTGATAAAGAACTCTTAGCGGTATGCAAAAGCT	750
Qy	181	GlyValAlaAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	751	GGAGCAGTAGCTGAAGAGGTCTGGCAGCAATTAGNACTGTGATTCATTTGGAGACAA	810
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	811	AAGAAGAACTTTGAAAGGTACACAAATAATTTAGAAGAGCTTAAAGAAATTTGGATAAAG	870
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	871	AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCATCTTAT	930
Qy	241	GlyLeuAlaPheTyrThrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	931	GCTCTGGCCCTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCATT	984
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	985	CGACAAGTACTCACTGATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTCTGGACAGGCA	1044
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	1045	TTCCCAAGCATTTGCAAGCATTTGCAAAATGCAAGGAGCAGCTTATGAAATCTTCAAGATA	1104
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1105	ATGATAATAAGCAAGTATTGACAGCTATTGAGAGGTGGGCACAAACAGATATATT	1164
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	1165	AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAAGTTAAG	1224
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1225	ATCTTGAAGGGCTGAACCTGAAGGTGCAAGGTGGGCAGACGGTGGCCCTGGTGGTGAAC	1284
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp	380
Db	1285	AGTGGCTGGGGAAGAGCACACAGTCCAGCTGTATGACAGGCTCTATGACCCACAGAG	1344
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	1345	GGAGTGTGAGTGTGATGACAGGATATTAGGACCATTAATGTAAGTTTCTTAACGGAA	1404
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	420
Db	1405	ATCATTTGGTGTGAGTCAAGAACCTGTATTGTTTGGCCACCACTAGCTGTAACACATT	1464
Qy	421	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1465	CGCTATGGCCGTAATATGTCCATGGATGAGATGAGAAAGCTGTCAAGGAAGCAAT	1524
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly	460
Db	1525	GCCTATGACTTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGCTGGAGAGAGGG	1584
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1585	GCCAGTTGAGTGGTGGCAGAGCAGAGGATCGCCATTTGACAGTCCCTGCTGTCACAC	1644
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1645	CCCAAGATCTCTGCTGGTGGCCAGCTGACGCTTGGACACAGAAAGCGAGAGTG	1704
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1705	GTTTCAGTGGCTGTGATAGGCCAGAAAGGTGGACCAACCATTTGTGTAGTGTATCGT	1764
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540

Db 1765 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGAGTCATGTTGGAG 1824  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
Db 1825 AAGGAAATCATGATGACTCATGAAGAAAGAGGATTTACTTCAAACTTGTCACAATG 1884  
Qy 561 Gln-----AspIleLysLysAlaAspGlnMetGluSerMetThr 574  
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAATGCGAGTCGATGAATCCAAAGAGTGAAT 1944  
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 599  
Db 1945 GATGCGCTTGAAATGCTTCAAAATTCAGATCCAGTCTCAATTAAGAAAGATCAACT 2004  
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607  
Db 2005 CGTAGGAGTGTCCGTGGATCAAGGCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
Qy 608 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysPProGluTTP 627  
Db 2065 GATGAAAGTATACCTCCAGTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATG 2124  
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
Db 2125 CCTATTATTTGCTGTGTTTCTGCTATTTATTAATGGAGCGCTGCAACAGCATTT 2184  
Qy 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 2185 GCATATATATTTCAAGATTATAGGGTTTTTACAGAAATGATGATCTGAAACAAAA 2244  
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
Db 2245 CGACAGAAATAGTAATGTTTCTACTATTGTTCTAGCCCTTGGAAATTTCTTTTATT 2304  
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706  
Db 2305 ACATTTTCTCCAGGTTTTTCATATTGGCAAAAGCTGGAGAGATCTCTCAACAGCGGCT 2364  
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLysGlu 726  
Db 2365 CGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTTGTGATGACCCTAAA 2424  
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
Db 2425 AACACCACTGGAGCATGTACTACAGCTCGCCAATGATGCTCAAGTTAAAGGGCT 2484  
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
Db 2485 ATAGTTTCCAGGCTTGCTGTAATTTACCCAGATATATAGCAAAATCTTGGGACAGGAAT 2544  
Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786  
Db 2545 ATATCTTTCATCTATGTTGGCACTAACACTGTGTACTCTTAGCAATTTGACCCATCAT 2604  
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
Db 2605 GCATAGCAGGAGTGTGAAATGAAATCTGTCTGCAACAGCACTGAAAGATAAGAA 2664  
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
Db 2665 GAACAGAGGTCGTGGAAAGATCGCTACTGAAAGCAATAGAAAACCTTCCGACCGTTGTT 2724  
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846  
Db 2725 TCITTGACTCAGGACGAGAAAGTTTGAACATATATGATGCTCAGAGTTTGCAGGTACCATAC 2784  
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866  
Db 2785 AGAACTCTTTGAGAAAGCACACATCTTTTGAATATACATTTTCTTCCACCCAGCAATG 2844  
Qy 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleuIleGlnAlaGlyArg 886  
Db 2845 ATGTATTTTCTTCTATGCTGATGTTTCCGGTTTGGAGCCTTACTTGTGGGCACATAAATC 2904

Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
Db 2905 ATGAGCTTTGAGGATGTTCTGTAGTATTATTTTTCAGCTGTGTTCTTGTGTCCTGCGGTG 2964  
Qy 907 GlyIleThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaHisLeu 926  
Db 2965 GGCGCAGTCAGTTTCTGCTGCTATGCAAGCCAAATATATCAGCAGCCACATC 3024  
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro 946  
Db 3025 ATCATGATCATTTGAAACAAACCCCTTTGATTGACGCTACAGCAGGAAGCCCTAATGCCG 3084  
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
Db 3085 AACACATTCGAAAGAAATGTCATATTGCTGAGTTGTTTCACTATCCACCCGACCG 3144  
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986  
Db 3145 GACATCCAGTGTCTCAGGCACTGAGCTGAGGTGAAGAGCACATGCTCTGAGCGGTTCTAGAC 3204  
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006  
Db 3205 GTGGCAGCAGTGGCTGTGGAGAGACACATGCTCCAGCTCTGAGCGGTTCTAGAC 3264  
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAlaLysAlaLysGluLeuAsnValGlnTTP 1026  
Db 3265 CCCTTGGCAGGAAAGTGTCTGCTGCAAGAAATAAAGCAGCTGAATGTTCACTGG 3324  
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046  
Db 3325 CTCGAGACACATCGGCATCTGCTCCAGAGGCCATCTCTGTTGACTGCAGCATTTGCT 3384  
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluLeuLysGluAla 1066  
Db 3385 GAGAACATTCCTATGAGACACACAGCGGGTGTGTGCACAGAAAGATCTGTGAGGGCA 3444  
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGlyGlyLeuProGluLysTyrAsnThrGln 1086  
Db 3445 GCAAAGGAGGCCAACATACATCGCTTCATCGAGTCACTGCTTAATAAATATAGCACTAAA 3504  
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106  
Db 3505 GTAGAGACAAAGGAACCTCAGCTCTCTGTGTGGCCAGAAACAACGCAATTCCTAGCTCGT 3564  
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 3565 GCCTTGTGTAGACAGCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624  
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
Db 3625 GAAAGTGAAGAGTTGTCCAGAGCCCTGGACAAAGCCAGAGAGCCGCGCCTGCAAT 3684  
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166  
Db 3685 GTGATTGCTCAGCGCTGTCCACCATCCAGATTCAGACTTAATAGTGTGTTCAGAAAT 3744  
Qy 1167 GlyLysIleLysGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186  
Db 3745 GGCAGAGTCAAGAGCATGGCAGCATCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804  
Qy 1187 LysLeuValAsnAlaGlnSer 1193  
Db 3805 TCAATGCTCAGTCTCCAGCT 3825  
RESULT 5  
ABA94365  
ID ABA94365 standard; DNA; 3860 BP.  
XX  
AC ABA94365;  
XX 26-MAR-2002 (first entry)  
DT  
XX

DE Human BCRP DNA related seq Id No. 1.

XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;  
KW hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;  
KW cardiant; gene therapy; ds.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 1..3843  
FT /\*tag= a

XX WO200192877-A2.

XX PD 06-DEC-2001.

XX PF 30-MAY-2001; 2001WO-US17459.

XX PR 31-MAY-2000; 2000US-0584586.

XX PR 29-MAY-2001; 2001US-0866866.

XX PA (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX PI Sorrentino B, Schuetz J;

XX DR WPI; 2002-114368/15.

XX DR P-PSDB; ABB07266.

XX PT Identifying a stem cell, for treating e.g., muscular dystrophy,  
PT myocardial infarction, Parkinson's disease, or neurodegenerative  
PT disorders, comprises detecting the expression of an ATP transport  
PT protein (BCRP) by a cell -

XX PS Disclosure; Page 53-55; 87pp; English.

XX CC The invention provides a method of identifying and/or isolating a stem  
CC cell that involves detecting the expression of an ATP transport protein  
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample  
CC comprising stem cells. The isolated stem cells may be used in the  
CC treatment of diseases such as muscular dystrophy, degenerative liver  
CC disorder, myocardial infarction, Parkinson's disease, degenerative  
CC disorders of the brain, and for tissue regeneration or replacement.  
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for  
CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
CC also be used as cell targets in gene therapy protocols. The present  
CC sequence represents a sequence related to the BCRP for which no relevant  
CC information has been provided in the specification.

XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	3860
Score:	3554.00	Matches:	692
Percent Similarity:	75.63%	Conservative:	246
Best Local Similarity:	55.58%	Mismatches:	263
Query Match:	59.12%	Indels:	36
DB:	24	Gaps:	9

US-09-873-409-6 (1-1195) x ABA94365 (1-3860)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeu 20

Db 151 ATGGTGGTGGGAACCTTGGCTGCATCATCCATGGGGCTGGACTTCTCATGATGCTG 210

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAAGATCTGATGTC 270

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43

Db 271 AACATCACTAATAGAAAGTATGATCAATGATACAGGGGTCTTCTATGAATCTGGAGGAAGAC 330

Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValaAlaIlePheGlyTyr 61  
Db 331 ATGACACAGATATGCTATATTTACAGTGGAAATGTGCTGGGGTCTGCTGCTCTTAC 390  
Qy 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 391 ATTGAGTTTCAATTTGGTGGCTGGCAGCTGGAACAAATACACAAATTTAGAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGCTATAATGCGACAGAGATAGCTGGTTGATGTGCACGATGTGGGGAG 510  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 511 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAATGAAGGTATTGGTGACAAAT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTACAGT 630  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GGTGGAAAGCTAACCTTGTGATTTGGCCATCAGTCTCTGTTCTTGGACTGTGACGTGCT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 691 GTCTGGCAAGATATCTATCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750  
Qy 181 GlyAlaValaIleGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGAATGATGATTCATTTGGAGACAA 810  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 811 AAGAAGAAGCTTGAAGGTACAAACAAATTTAGAAGAGCTAAGAAGATTGGATAAG 870  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 871 AAAGCTATTACAGCAATATTCTATAGTGTGCTGCTCTGCTGCTATGATCATCTTAT 930  
Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 931 GCTTGGCTCTTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 984  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 985 GCACAGTACTCAGTGTATCTTCTGTATTAATGGGGCTTTAGTCTGGACAGGCA 1044  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1045 TCTCCAAGCATTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104  
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1105 ATTGATAATAAGCCAAAGTATTGACAGCTATTGCAAGAGTGGGCAACAAACAGATAAT 1164  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1165 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAGTTAG 1224  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyThrValAlaLeuValGlyLeu 360  
Db 1225 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGACAGCGTGGCCCTGTTGGAAAC 1284  
Qy 361 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1285 AGTGCGTGTGGGAAGAGCACACAGTCAGCTGATGACAGAGGCTCTATGACCCACAGAG 1344  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1345 GGGATGGTGTGATGGAGCAGGATATTAGACCATTAATGTAAAGTTTCTACGGGAA 1404  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420











```
Db 3155 CCACCTGACACAGATCCCGCTGCTCCAGGGCTGAGCTCGAGTGAGAGGCCAG 3214
Qy 983 ThrValAlaPheValGlySerGlyCysGlyLysSerThrSerValGlnLeuGln 1002
Db 3215 ACCTGGCCCTCGTAGTAGCAGTGCTGGGAAGACACAGTGTCTCAGCTCTAGAG 3274
Qy 1003 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3275 CGCTTCTATGACCCCTGGCTGCTCAGTGCTAAATTGATGGCAAGAGATAAAGCACCTG 3334
Qy 1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db 3335 AATGTCAGTGCTCGAGCACCTGGGCATCGTCTCAGAGGCCATCTGTTGAC 3394
Qy 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3395 TGCAGATTGCCGAGAACATTGCCATGAGACACAGCCGGTCTGATCACATGAAGAG 3454
Qy 1063 IleLysGluAlaAlaAsnAlaAlaIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3455 ATTATGCGAGCGACCAAGGAGGCCAACATACACCACTTTCATCGACACACTCCCTGAGAA 3514
Qy 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3515 TACAACACAGAGTAGGACAAAGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT 3574
Qy 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db 3575 GCCATAGCTCGGCTCTGTTAGACAGCTCATATTTGCTTTGGTGAAGCTACATCA 3634
Qy 1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3635 GCTCTGGATACAGAAAGTCAAAAGGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAAGC 3694
Qy 1143 ArgThrCysLeuValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3695 CGCAGCTGATGTGATGCCCGCCAGCTGTCACCATCCAGATTCAGATTAAATAGTG 3754
Qy 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db 3755 GTGTTTCAGATGGCAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGCCAGAAA 3814
Qy 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3815 GCATCTATTTTCCATGCTCAGTGCCAGGCT 3847

RESULT 7
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.
XX
AC AAD03505;
XX
DT 13-JUN-2001 (first entry)
DE
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FH CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT allele replace (91, T)
FT /*tag= b
FT allele replace (607, C)
FT /*tag= c
XX
PN W0200123540-A2.
```

```
XX 05-APR-2001.
PD
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00309.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX PT for determining the bioavailability of drugs and for screening for dog
XX PT PGP inhibitors -
XX
XX Claim 9; Page 93-99; 11lpp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX CC as multidrug transporter (MDRI) and nucleic acids encoding them.
XX CC The invention also includes fragments and biologically functional
XX CC variants of dog P-glycoprotein. PGP and their nucleic acids are
XX CC useful for determining the bioavailability of drugs and for
XX CC screening PGP inhibitors. They are useful for the diagnosis and
XX CC treatment of conditions characterised by PGP activity, by
XX CC reducing or increasing PGP activity in a cell. PGP nucleic acids
XX CC are used as oligonucleotide probes. Complements of PGP nucleic
XX CC acids are useful as antisense oligonucleotides, to induce a PGP
XX CC 'knockout' phenotype. They are used to prepare a non-human
XX CC transgenic animal, which are valuable as genetic models for
XX CC human diseases.
XX CC The present sequence is dog P-glycoprotein (PGP) allelic variant
XX CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
XX CC exporting small molecules across the cell membrane. This enzyme
XX CC is a member of the ABC transporter family.
XX
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4279
Score: 3554.00 Matches: 682
Percent Similarity: 74.74% Conservative: 238
Best Local Similarity: 55.40% Mismatches: 269
Query Match: 59.12% Indels: 42
DB: 22 Gaps: 8

US-09-873-409-6 (1-1195) x AAD03505 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValaenGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAACTGCTGCCATCATCCATGAGGTGCACTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CCAGTTATAATTAATGAAGTATTACGAAACATACCAACATTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCAACGATGCTGCTATTATTACAGTGGGATCGGTGCTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTTCATTCTGTCCTGCGCAGGAGACAGATCTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AAACAATTTTTTCATGCTATCATGCGAGGAGATTTGGCTGTTTGCATGACGCTT 526
```



[illegible][illegible]

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:

Query Match: 59.03% Indels: 42  
DB: 22 Gaps: 8  
US-09-873-409-6 (1-1195) x AAD03488 (1-4279)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 167 ATGTGGTGGGCAATGGCTGGCATCATCGAGCTCACTCCCTCTCATGATGCTG 226  
QY 21 ValIleuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
DB 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGAAATTCAGAAACAAAACTTTT 286  
QY 35 -----ValGlnThrAsnThrThrThrThrThrThrPheArg----- 44  
DB 287 CCAGTTAATTAATGAAGATTAATACGAACAATACACAACATTTTCATCAACCATCTGGAG 346  
QY 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePhe 59  
DB 347 GAGGAATGACCATGCTATGCTTATTACAGTGGATCGGTGCTGGCTGCTGCT 406  
QY 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79  
DB 407 GCTTACATCCAGGTTTCATCTGCTGCTGCGACGAGAGACAGATACATAAAATTAGA 466  
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
DB 467 AACAAATTTTTCATGCTATCATGCGACAGAGATTTGGCTGGTTTGACGTCGATGAGTT 526  
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
DB 527 GGGGAGCTTAACACCGGCTCACAGACGATGCTCCAAAATCAATGAAGAAATTGGCGAC 586  
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
DB 587 AAATTTGGAATGTTCTTCTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646  
QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
DB 647 ACAGTGGTTGGAAGCTAACCCCTTGATTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706  
QY 159 AlaAlaLysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
DB 707 GCGCCATCTGGGCAAGATACATCTTCACTTACTATAAAGAACTCTTGGCGCTATGCA 766  
QY 179 LysAlaGlyAlaValAlaGluGlnValLeuSerSerIleArgThrValIleAlaPheArg 198  
DB 767 AAAGCTGGAGCGAGTAGCTGAAGAAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 826  
QY 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
DB 827 GGACAAAGAAAGAACTTGAAGGTACACAAATAATTTAGNAGAGTAAGGAAATTTGGG 886  
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238  
DB 887 ATAAAGAAAGCTATCACGGCCCAACATTTCTATTGTCGCGCTTCTTATTGATCATGCA 946  
QY 239 ThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258  
DB 947 TCATATGCTCTGGCTTTCTGGTATGGACCTCTTGTGCTCTCTCCAGTGAA-----TAT 1000  
QY 259 ThrIleGlyThrValLeuAlaValPheSerValIleHisSerTyrCysIleGly 278  
DB 1001 TCTATTGGACNAGTACTACTGCTCTTCTTCTGATTAATTTGGGGCTTTTAGTATTGGA 1060  
QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298  
DB 1061 CAGGCATCCCAACCACTTGAAGCATTTTGCAACCGCAAGAGAGAGCAGCTTATGAAATCTTC 1120  
QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
DB 1121 AAGATAATTGCAATAACCAAGCATTTGACAGCTATTTCGAAGAGTGGACATAAACCCAGAT 1180

QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
DB 1181 AATATTAGGAAATTTGGATTTCAAAATTTGTTCACTTCAGTTACCTTTCTCGAAAGAA 1240  
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358  
DB 1241 GTTAAGATCTTAAGGGTCTCAACCTGAAGGTTCAAGTGGGAGACAGTGGCGCTGGTT 1300  
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378  
DB 1301 GGGAAACAGTGGCTGGGGAAGAGACACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360  
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398  
DB 1361 ACAGATGGCATGCTGTTATGATGACAGACATTTAGGACCATAAATGTAAGGCATCTT 1420  
QY 399 ArgAspHisIleGlyValValSerGlnGluProValIlePheGlyThrThrIleSerAsn 418  
DB 1421 CGGAAATTTACTGGTGTGTGAGTCAGGAGCCTGTTGTTTGGCCACGATAGCTGAA 1480  
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438  
DB 1481 AACATTCGTATGGCCGCGCAAAATGTCACCATGGATGAGATTGAGAAGCTGTTAAGAA 1540  
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458  
DB 1541 GCCAATGCCTATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTCTGGTTGGAG 1600  
QY 459 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478  
DB 1601 AGAGGGCCCGAGCTGAGTGGTGGACAGAAACAGAAATCGCCATTCCTCGGGCCCTGGTT 1660  
QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498  
DB 1661 CGCAACCCCAAGATCTTCTGCTGATGAGGCAACCTCAGCTCTGGACACTGAAGTGAA 1720  
QY 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518  
DB 1721 GCAGTGGTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1780  
QY 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538  
DB 1781 CATCGTTGTGTACATGTTGATGCGGATGTCATTCGCTGGTTTGTGATGATGAGTCAIT 1840  
QY 539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558  
DB 1841 GTGGAGAAAGAAATCATGATGAACTCATGAAAGAGAGAGGCAATTTACTTCAAACTGTC 1900  
QY 559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578  
DB 1901 ACAATGCAG---ACAAGAGGAAATGAATAGTTAGTAAATAATGCCACTGTGTAATCCAA 1957  
QY 579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597  
DB 1958 AGTGAAGTATGCTCTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014  
QY 598 -----AspLysAlaGluGluSer 603  
DB 2015 AGAAGATCACTCGCAGGAGTATACATGCACCAAGCCCAAGCAAGCAAGCTGGTACA 2074  
QY 604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsn 623  
DB 2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCTTCTGGAGGATTTCTGAAGCTGAAC 2134  
QY 624 LysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643  
DB 2135 TCAACTGAATGGCTTATTTTGTGGTGGTATATTGTTGCTATTATAACGAGGCGCTG 2194  
QY 644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662  
DB 2195 CAACCGACATTTTCAATATATATTTTCAAGGATATAGGATCTTTACCCGAGATGAGAT 2254  
QY 663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682

```
Db 2255 CCTGAACAAACCGACAGAAATAGTAAATGTTTCTGTATGTTTCTAGTCTCTGGAATT 2314
Qy 683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db 2315 ATTTCTTTTATACATTTTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCTTC 2374
Qy 703 ThrMetArgLeuArgHisLeuAlaPheTysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db 2375 ACTAAGCGGCTTCGATACATACATGTTTTCAGATCCATCTGAGACAGGATGTCAGCTGGTT 2434
Qy 723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db 2435 GATGACCCCTAAAAACACCACTGGAGCATTCACAAACAGGCTTGCCCAATGATCGGCTCAA 2494
Qy 743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db 2495 GTTAAAGGGGCTATAGTTCCAGGCTTGCTGTCATTTACCCAGATATAGCAATCTTGGG 2554
Qy 763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db 2555 ACAGGCAATATTATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614
Qy 783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db 2615 GTACCCATCATTTGCAATAGCAGGAGTTGTGAATGAAATGTTGTCTGGCAAGCACTG 2674
Qy 803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlnAlaLeuGluAsnIle 822
Db 2675 AAGATTAAGAAGAGCTAGAGAGCTGGGAAGATGTCTACAGAACCAATCGAAACTTC 2734
Qy 823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db 2735 CGAAGCTGTTGTTCTTTGACTCGGAGCAGAGTTTGAATACATGATGCACAGAGTTTG 2794
Qy 843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPhe 862
Db 2795 CAAGTACCATACAGAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTCTATC 2854
Qy 863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db 2855 ACCAGCGCAATGATGATTTTCTCTATGCTGGCTGTTCCGGTTTTGGTCTACTTGGTG 2914
Qy 883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db 2915 GCATAAGTTCATGACATTCATGAGTGTCTTTTGGTATCTCAGCTATTGCTTTGGT 2974
Qy 903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db 2975 GCCATGGCAGTGGCGAGGTCACTTCAATTTGCTCTGACTATGCCAAAGCCAAAGTATCA 3034
Qy 923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db 3035 GCAGCCAGCTCATCATGATCATTTGAAAGAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094
Qy 943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPheTyr 962
Db 3095 GGCCTCAGCCAAATACGTTGGAGGAATGTGCATTTAATGAGGTGCGTTTCACTAT 3154
Qy 963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db 3155 CCCACTCGACCACACATCCCCTGCTCCAGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214
Qy 983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db 3215 ACGCTGCGCCCTCGTAGTACAGTGGCTGGGAAGACGACAGTTGTTTCAGCTCCTAGAG 3274
Qy 1003 ArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db 3275 CGTTCTATGACCCCTTGGCTGTTGCTGCTAATGTATGGCAAGAGATAAGACACCTG 3334
Qy 1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
```

```
Db 3335 AATGTCAGTGGCTCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCTGTTGAC 3394
Qy 1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db 3395 TGCAGCATTTGCCGAGAACATTGCTTATGAGACAAACAGCCGGTCTGTATCACATGAAGAG 3454
Qy 1063 IleLysGluAlaAlaAsnAlaIleHisSerPheIleGluGlyLeuProGluLys 1082
Db 3455 ATTATGCGAGGCCAACGAGGCCAACATACACCTTTCATCGAGACACTCCCTTGAGAA 3514
Qy 1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db 3515 TAAACACCGAGTAGGACAAAGAACCCAGCTCTCTGTGGCCAGAAACAGCCGAT 3574
Qy 1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSer 1122
Db 3575 GCCATAGCTCGGCTCTTTGTAGACAGCCTCATATTTTGTGATGAAGCTACATCA 3634
Qy 1123 AlaLeuAspAsnAspSerGlnLysValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db 3635 GCTCTGGATACAGAAAGTGAAGGTTGTCCAAGAGCCCTCGACAAAGCCAGAGAGC 3694
Qy 1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db 3695 CCGACTGCTATGTGATCGCCACCGCTTGTCCACATCCAGATGCAGATTTAATAGTG 3754
Qy 1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db 3755 GTGTTTCAGATGGCAAGTCAAGGAGCATGCGACACATCAACAGCTGCTGGCCAGAAA 3814
Qy 1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db 3815 GGCATCTATTTTCCATGTCAGTGTCCAGGCT 3847

RESULT 9
AAD03489
ID AAD03489 standard; cDNA; 4317 BP.
XX
AC AAD03489;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #2.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 70..3912
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #2"
XX
PN WO200123540-A2.
XX
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI; 2001-235373/24.
XX
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
```

Claim 1; Page 66-72; 11lpp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterized by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) cDNA. The  
 CC PGP enzyme functions as an efflux pump exporting small molecules  
 CC across the cell membrane. This enzyme is a member of the ABC  
 CC transporter family.

XX Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

# Alignment Scores:

Pred. No.: 0 Length: 4317  
 Score: 3549.00 Matches: 680  
 Percent Similarity: 74.65% Conservative: 239  
 Best Local Similarity: 55.24% Mismatches: 270  
 Query Match: 59.03% Indels: 42  
 DB: 22 Gaps: 8

US-09-873-409-6 (1-1195) x AAD03489 (1-4317)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 217 ATGTGGTGGGCAATAGGGCTGCCATCATCCATGGAGCTGCACCTCTCATGATGCTG 276  
 Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
 Db 277 GTTTTGGAAACATGACAGATAGCTTTGCAATGTCAGGAATTTCAAGAAACAAACTTTT 336  
 Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
 Db 337 CCAGTTATAATTAATGAAGTATTACCAACAATACACACATTTTCAACACCATCTGGAG 396  
 Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuPhe 59  
 Db 397 GAGAAATACCCAGTATGCTATTATACAGTGGATCGGTGCTGCGCTGGTGGCT 456  
 Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
 Db 457 GCTTACATCCAGGTTTCTTCTGTGCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516  
 Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99  
 Db 517 AAACAATTTTTCATGCTATCATCCGACAGAGATTGGCTGGTTGACGTGCATGACGTT 576  
 Qy 100 GlyGluLeuAsnThrArgMetThr--AspIleAspIleSerAspGlyIleGlyAsp 118  
 Db 577 GGGGAGCTTAACACCCGGCTCACAGACGATGCTCTCAAAATCAATGAAGAAATGGCGAC 636  
 Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
 Db 637 AAAGTTGGAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT 696  
 Qy 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
 Db 697 ACACCTGGTTGGAAGCTAAACCTTGTGATTTGGCCATCAGCCCTGTTCTTGGCACTTCA 756  
 Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
 Db 757 GCGCCCATCTGGGCAAGATCATCTTCTTACTGATTAAGAATCTTTGGCCCTATGCA 816  
 Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198

Db 817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 876  
 Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
 Db 877 GGAACAAAGAAAGAACTTGAAAGGTACAAACAAATTTAGAGAGACTAAAGAAATGGG 936  
 Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyValaValTyrPhePheMetAsnGly 238  
 Db 937 ATAAAGAAAGCTATCACGGCCCAACATTTCTATGGTGGCGCTTCTTATGAICTATGCA 996  
 Qy 239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluPheGlyTyr 258  
 Db 997 TCATATGCTCTGGCTTCTGGTATGGGACCTCTTGGTCTCTCCAGTGAA-----TAT 1050  
 Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278  
 Db 1051 ACTATTGGGACAGGTACTCACTGCTTCTTCTTCTTATTAATGGGGCTTTTAGTATGGA 1110  
 Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298  
 Db 1111 CAGGCATCCCAAGCATTGAAGCATTTGCAACGCAAGAGGAGCAGCTTATGAATCTTC 1170  
 Qy 299 GlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
 Db 1171 AAGATAATTGACAATAAACCAAGCATTTGACAGCTATTTCGAAGAGTGCACATAACAGAT 1230  
 Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
 Db 1231 AATATTAAAGGAAATTTGGAATTTCAAAATGTTCACTTCAGTTACCCCTTCTCGAAAGAA 1290  
 Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358  
 Db 1291 GTTAAGATCTTAAGGGTCTCAACCTGAAGTTTCAGGTGGGAGAGTGGCGCTGGTT 1350  
 Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378  
 Db 1351 GGGACAGTGGCTGCGGGAAGAGACACGCTGCAGCTGATGCAGAGGCTCTATGACCCC 1410  
 Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398  
 Db 1411 ACAGATGGCATGCTGCTGTTATGATGGACAGGACATTAGGACCATTAATGTAAGCATCTT 1470  
 Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418  
 Db 1471 CGGAAATTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1530  
 Qy 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438  
 Db 1531 AACATTTCGCTATGGCCGCAAAATGTCAATGGATGAGATTGAGAAAGCTGTAAGGAA 1590  
 Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458  
 Db 1591 GCCAATGCCCTATGATTTTATCATGAAACTACCTTAATAAATTTGACACTCTGTTGGAGAG 1650  
 Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478  
 Db 1651 AGAGGGCCCGCTGAGTGGTGGACAGAAACAGAGAAATGCCCATTTGCTCGGGCCCTGGTT 1710  
 Qy 479 ArgAsnProLysIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 498  
 Db 1711 CGCAACCCCAAGATTCTTCTGCTGGATGAGCAACCGTCAGCTCTGGACACTGAAAGTAA 1770  
 Qy 499 SerAlaValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518  
 Db 1771 CGAGTGGTTCAGTGGCCCTGGATAGGCGAGAAAGGCGGAGCTACCATTTGTGAGTCT 1830  
 Qy 519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538  
 Db 1831 CATCGTTTGTCTACAGTTCGTATGCGGATGTCATTCCTGCTGCTGTTTGTGATGAGTCA 1890  
 Qy 539 AlaGluGlyAlaHisAlaGluLeuMetAlaLeuArgGlyLeuTyrTyrSerLeuVal 558



Db	1891	GTGAGAGAAAGGAATCATGATGAAGCTCATGAAGAGAGAGGGCATTTTACTTCCAAACTTGTCT	1950
Qy	559	MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu	578
Db	1951	ACAATGCAG--ACAAGAGGAATGAAATTGAGTTAGAAAATGCCACTGGTGATCCAAA	2007
Qy	579	ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle---	597
Db	2008	AGTGAAGAGTGAATCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA	2064
Qy	598	-----AspLysAlaGluGluSer	603
Db	2065	AGAAGATCAACTCGCAGGAGTATACATGCACCACCAAGGCCAAGACCAAGAACTGGTGA	2124
Qy	604	ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn	623
Db	2125	AAAGAGGACTTGAATGAGAATGTTCTTCAGTTTCTTCTGGAGGATTCGAAGCTGAAC	2184
Qy	624	LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal	643
Db	2185	TCAACTGAATGGGCTTATTTTGTGGTGGTATATTTTGTGCTATTATATAACGAGGCGCTG	2244
Qy	644	HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe--GlyAsnAsnAsp	662
Db	2245	CAACCAGCATTTTTCAATAATATTTTTCAAGGATATATAGGATCTTTTACCCGAGATGAGAT	2304
Qy	663	LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal	682
Db	2305	CCTGAACAACAAACACAGAAATAGTACATGTTTTCTGTATTGTTCTTAGTCTTTGGAA	2364
Qy	683	IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu	702
Db	2365	ATTTCTTTTATTACATTTTCTCCAGGCTTCACATTTGGCAAGCTGGGAGATCCTC	2424
Qy	703	ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe	722
Db	2425	ACTAAGCGGCTTCGATACATGTTTTTCAGATCCATGCTGAGACAGATGTCCAGCTG	2484
Qy	723	AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln	742
Db	2485	GATGACCTTAAAAACACCCTGGAGCATTCACACAGCGCTTGCAATGATCGCGCTCAA	2544
Qy	743	IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly	762
Db	2545	GTTAAGAGGGCTATAGTTCACGGCTTGCTGTCAATTCACCAGAAATATAGCAAACTCTT	2604
Qy	763	LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle	782
Db	2605	ACAGGCATTTATATCTCTAAATCTATGTTGGCAATTAACACTTTTACTCTTTAGCA	2664
Qy	783	AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn	802
Db	2665	GTACCCATCATTCGAATAGCAGGAGTTGTGAATGAATAATGTTGCTGGACAAGCACTG	2724
Qy	803	LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle	822
Db	2725	AAAGATGAAGAGGCTAGAAGAGCTGGGAGATTGCTACAGAGCCATCGAAAACCTC	2784
Qy	823	ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu	842
Db	2785	CGAACTGTGTCTTTTGAATCTCGGAGACAGAAGTTTTGAATACATGTATATGCACAG	2844
Qy	843	GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPhe	862
Db	2845	CAAGTACCATACAGAAACTCTTTGAGGAAGACACATCTTCGGGGCTTCATTTTCTATC	2904
Qy	863	SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle	882
Db	2905	ACCCAGGCAATGATGATTTTTTCTATGCTGGCTGTTTCCGGTTGGTGCCCTACTTGG	2964
Qy	883	GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly	902
Db	2965	GCAATGAGTTCATGAACCTTCAGGATGTTCTTTGGTATTCTTCAGCTATTGCTTTGG	3024

Qy	903	AlaValAlaIleGlyIysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly	922
Db	3025	GCCATGCGAGTGGGGCGAGTCAGTTCAITTTGCTCTGACTATGCCAAAGCCAAAGATATCA	3084
Qy	923	AlaAlaHisLeuPheAlaLeuLeuGluLysProAsnIleAspSerArgSerGlnGlu	942
Db	3085	GCAGCCCACTCATCATGATCATTTGAAAAGGCCCTCGATTGACAGCTACAGCGCCCTCAC	3144
Qy	943	GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr	962
Db	3145	GGCTCTCAAGCCAATACGTTGGNAGGAATGTGACATTAATGAGTCTGTTCAACTAT	3204
Qy	963	ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys	982
Db	3205	CCCACTCGACAGACATCCCGGTGCTCCAGGGGCTGAGGCTCGAGGTGAAGAAGGGCCAG	3264
Qy	983	ThrValAlaPheValGlySerSerGlyCysGlyIysSerThrSerValGlnLeuLeuGln	1002
Db	3265	ACGCTGGCCCTCGTAGTATGACATGGCTGTGGGAAGACACAGTTGTTCACTCTCTAGAG	3324
Qy	1003	ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu	1022
Db	3325	CGCTTCTATGACCCCTGGCTGGTTCAGTGTCAATTGATGGCAAGAGATAAAGCACCTG	3384
Qy	1023	AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn	1042
Db	3385	AATGTCAGTGGTCCGAGACACCTGGGCATCGTCTCAGGAGCCCATCTCTGTTTAC	3444
Qy	1043	CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu	1062
Db	3445	TGCAGCATTCGCGAGAACATTCGCTATGAGACAACAGCGGGTCTGATCATCATGAAGAG	3504
Qy	1063	IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys	1082
Db	3505	ATTATGCGAGCGACCAAGGAGGCAACATACACCACTTCATCGACAGACTCCCTCGAGAAA	3564
Qy	1083	TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu	1102
Db	3565	TACAACACAGATGAGAGACAAGAAGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCAT	3624
Qy	1103	AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer	1122
Db	3625	GCCATAGCTCGGCTCTGTTAGACAGCTCATATTTTGGTCTTTGGATGAAGTACATCA	3684
Qy	1123	AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly	1142
Db	3685	GCTCTGGATACAGAAAGTGAAGAGTGTGTCGAAGAGCCCTGGACAAAGCCAGAGAGGC	3744
Qy	1143	ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal	1162
Db	3745	GGCACTTGCATTGTGATPGCCCAACCCCTGTCCACCATCCAGATGCGAGATTTAATAGTG	3804
Qy	1163	ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuArgAsnArg	1182
Db	3805	GTGTTTCAGATGCGAAAGTCAAGGAGCATGCGACACATCAACAGCTGCTGCTCAGAAA	3864
Qy	1183	AspIleTyrPheLysLeuValAsnAlaGlnSer	1193
Db	3865	GGCATCTATTTTCCATGATCAGTGTCCAGGCT	3897



XX Macaca fascicularis.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 100..3942  
 FT /\*tag = a  
 FT /product= "PGP"  
 FT /note= "P-glycoprotein"  
 PN  
 XX WO200123565-A1.  
 XX  
 PD 05-APR-2001.  
 XX  
 XX 28-SEP-2000; 2000WO-US26592.  
 XX  
 XX 28-SEP-1999; 99US-0156921.  
 PR 12-OCT-1999; 99US-0158818.  
 XX  
 XX (GENT-) GENTEST CORP.  
 XX  
 XX Stocker PJ, Steimel-Crespi DT, Crespi CL;  
 PI WPI; 2001-316136/33.  
 XX P-PSDB; AAB81064.  
 DR  
 DR  
 XX Novel isolated nucleic acid encoding cynomolgus monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 PT cell -  
 XX  
 XX Example 1; Page 51-57; 84pp; English.  
 PS  
 XX This invention relates to a polynucleotide sequence encoding a  
 CC cynomolgus monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention  
 CC includes a cynomolgus monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents cDNA encoding  
 CC cynomolgus monkey P-glycoprotein.  
 XX  
 SQ Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:  
 Pred. No.: 0 Length: 4186  
 Score: 3548.00 Matches: 679  
 Percent Similarity: 75.37% Conservative: 248  
 Best Local Similarity: 55.20% Mismatches: 261  
 Query Match: 59.02% Indels: 42  
 DB: 22 Gaps: 10

US-09-873-409-6 (1-1195) x AAF86127 (1-4186)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 250 ATGGTGGGAACTTGGTGGCCATCATCGAGAGCTGCGATTCCTCTCATGATGCTG 309  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33  
 DB 310 GTGTTGGACATGAGGATACCTTGGCAATGCAGGAATTTAGGAGATTTAGGAGCT 369  
 QY 34 LeuValGlnThrAsnThrTySer----- 41  
 DB 370 CTGTG---ACTAATAGCATTAATCACTGATACAGTGCCTGCATGAATCTGGAGGAA 426  
 QY 42 ---PhePheArgLeuThrLeuTyTyValGlyIleGlyValAlaAlaLeuIlePheGly 60  
 DB 427 GATATGACCAGGATGCGCTATTATTACAGTGAATTGGTCTGCGGGTGTGCTGTCT 486

QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
 DB 487 TACATTACAGTTTCAATTTGGTGCCTGGCAGCTGGAGACAAATACACAAAATTGAAAA 546  
 QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
 DB 547 CAGTTTTCATGCTATAATGCACAGGAGATAGGCTGGTTGATGTGCACGATGTGGG 606  
 QY 101 GluLeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLys 119  
 DB 607 GAGCTTAACACCCGGCTTACAGATGATCTCCAAGATTAAAGAAGAAATGGTGACAA 666  
 QY 120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139  
 DB 667 ATTGGAATGTTCTTTCAGTCAATGGCAATTTTCACTGGGTTTATAGTAGGATTACA 726  
 QY 140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159  
 DB 727 CGTGGTTGAAGCTAACCTTGTGATTTGGCCATCAGTCTCTTCTGGACTGTGAGCT 786  
 QY 160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTrpSerLys 179  
 DB 787 GCAGTCTGGGCAAGATACCTGCTTCATTACTGATAAAGAACTCTTAGCTTATGCAAAA 846  
 QY 180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199  
 DB 847 GCTGGAGCAGTAGCTGAGAGGCTTGGCAGCAATTAGAACTGTGATTGTCATTGGAGGA 906  
 QY 200 GlnGluLysGluLeuGlnArgTyThrGlnAsnLeuLysAspAlaLysAspPheGlyIle 219  
 DB 907 CAAAAGAAAAGACTCGAAAGGTACACAAAAATTTAGAAGAAGCTAAAGAAATTGGGATA 966  
 QY 220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyPheMetAsnGlyThr 239  
 DB 967 AGAAGACTATTACGCCAATATTTCTATAGGTGCTGCTTCTCTGCTATCTATGATCATCT 1026  
 QY 240 TyrGlyLeuAlaPheTrpTyGlyThrSerLeuIleLeuAsnGlyGluProGlyTyThr 259  
 DB 1027 TATGCTCTGGCTTCTGGTATGGACACCTTGGTCTCTCAAGGAA-----TATTCT 1080  
 QY 260 IleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrcysIleGlyAla 279  
 DB 1081 ATTGGCAAGTAGTACTACTGTATTTCTGTATTAAATGGGGCTTTTAGTGTGGACAG 1140  
 QY 280 AlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGln 299  
 DB 1141 GCATCTCAAGCATTTGAACATTTGCAATGCAAGGAGGAGCAGCTTTTGAATCTTCAAG 1200  
 QY 300 ValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrcysProGluSer 319  
 DB 1201 ATAATTGATAATAAGCAAGTATTGACAGTATTTCGAAGAGTGGGCACAAACCCAGATAAT 1260  
 QY 320 IleGluGlyThrValGluPheLysAsnValSerPheAsnTyrcysProSerArgProSerIle 339  
 DB 1261 ATTAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCTATCTCGAAAGAAGTT 1320  
 QY 340 LysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGly 359  
 DB 1321 AGATCTTGAGGGCTTGAACTGAGGTGCAGTGGCAGACGGTGGCCCTGGTTGGA 1380  
 QY 360 LeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrcysProAsp 379  
 DB 1381 AACAGCGGCTGTGGGAAGAGACACACGGTCCAGTCTGATGCAGAGGCTTTATGACCCACA 1440  
 QY 380 AspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrcys 399  
 DB 1441 GAGGGCATGCTGCTGATGTTGATGGACAGATATTAGACCATTAACCTAAGGTTTCTACCG 1500  
 QY 400 AspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsn 419  
 DB 1501 GAAATCATCGTGTGCTGAGTACGAACTGTATTGTTGCCACCATGATAGTGAAC 1560  
 QY 420 IleLysTyrcysArgAspValThrAspGluGluMetGluArgAlaAla-gluAla 439





Qy	241	GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAenGlyGluProGlyTyrThrIle	260
Db	931	GCTCTGGCCCTTCTGGTATGGACACCTTGGTCTCTCAGGGAA-----TATTCTATT	984
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	985	GGACAGTACTCACTGTTCTTTCTGTTATTAATGTTGGGCTTTTAGTGTGGACAGCA	1044
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGluVal	300
Db	1045	TCTCCAAGCATTGAAGCATTTCAAATGCAAGAGGAGAGCTTATGAATAATCTTCAGATA	1104
Qy	301	IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1105	ATTGATTAATAGCCAGTATTGCACGCTATTTCGAGAGTGGGCACAAACAGATAAATT	1164
Qy	321	GluGlyThrValGluPheLysAenValSerPheAenTyrProSerArgProSerIleLys	340
Db	1165	AAGGCAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGATTAA	1224
Qy	341	IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1225	ATCTTTGAAGGCCCTGAACTCGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTTGGAAC	1284
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	1285	AGTGGCTGTGGGAAGACACACAGTCCAGCTGTATGTCAGAGGCTCTATGACCCACAGAG	1344
Qy	381	GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyrArgAsp	400
Db	1345	GGGATGGTCAGTGTGATGCACAGGATATTAGGACCATAAATGTAAGCTTTCTACGGAA	1404
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	420
Db	1405	ATCATTTGGTGTGGTAGTCAGGAACCTGATTGTTGGCCACCAGCAGTAGCTGAAACCAT	1464
Qy	421	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1465	CGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAGAAAGCTGTACAGAAAGCCAT	1524
Qy	441	AlaTyrAspPheIleMetGluPheProAenLysPheAsnThrLeuValGlyLysGly	460
Db	1525	GCCTATGACTTTATCATGAACTGCCTCATAAATTTGACACCCCTGTTGGAGAGAGGG	1584
Qy	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1585	GCCAGTTGAGTGTGGCAGAGACAGAGATGCGCCATTGCACGTGCCCTGGTTGCGAAC	1644
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1645	CCCAGATCTCTCTGCTGGATGAGGCCACGCTTGGACAGAAAGCGAAGCGATGT	1704
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1705	GTTCCAGGTGGCTCTGGATAAGGCCAGAAAGGTTCGACCCACCATTTGTATAGCTCATCGT	1764
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1765	TGTCTACAGTTCTGTAATGCTGACATCGTCATCGTGGTTTCGATGATGGAGTCAATTGG	1824
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
Db	1825	AAAGGAATCATGATGAACTCATGAAGAAAGGACATTTACTTCAAACTTGTCCAAATG	1884
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	1885	CAGACAGCGAAATGAAGTTGAATTTAGAAATGACAGCTGATGAATCCAAAAGTGAAATT	1944
Qy	575	TyrSerThrGluArgLysThrAenSerLeuProLeuHisSerVal-----	589
Db	1945	GATGCTTGGAAATGCTTCAAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT	2004

Qy	590	---LysSerIleLys---SerAspPheIleAspLysAlaGluGluserThrGlnSerLys	607
Db	2005	CGTAGGAGTGTGCGTGGATCAAAAGCCCAAGACAGCTTAGTAGTACAAAGAGCGCTGTG	2064
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTirp	627
Db	2065	GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATATAGNAGCTAATTTTAACCTGAATGG	2124
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2125	CCTTATTTTGTGTGTGTATTTTGTGCCATTATAATAGGAGCGCTCAACACGACATTT	2184
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2185	GCAATAATATTTTCAAAGATTTATAGGGGTTTTACAGAAATTGATGATCCTCAACAAAAA	2244
Qy	667	LysHisAspAlaGluIleLysSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2245	CGACAGAATAGTAACCTGTGTTTACATATTGTTTCTAGCCCTTGGAAATATTCTTTTATT	2304
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2305	ACATTTTTCCTTCAGGTTTCACATTTGGCAAAAGCTGGAGAGATCTCTCACCAGGGCTC	2364
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirpPheAspGluLysGlu	726
Db	2365	CGATACATGTTTTCCGATTCATGCTCAGACAGGATGTAGTTGGTTGTATGATGCCCTAAA	2424
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2425	AACACCACTGGAGCATTTGACTACAGGCTCGCCAGATGATGCTCAAGTTAAAGGGCT	2484
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	2485	ATAGTTCAGGCTGCTGTATTAATCCACAGATATAGCAATCTTGGGACAGGAATAATT	2544
Qy	767	IleSerPheIleTyrGlyTirpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2545	ATATCCTTCATCTATGTTGGCAACTAAACACTGTTTACTCTTAGCAATTTGTACCCATCAT	2604
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	2605	GCAATAGCAGGAGTTGTGAATGAAAAATGTTGTCTGGCAAGCACTGAAAGATAGAA	2664
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	2665	GAACTAGAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTT	2724
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	2725	TCTTTTGACTCAGGAGCAGAAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCAC	2784
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	2785	AGAAACCTCTTTGAGGAAGACACATCTTTTGGAAATTTACATTTTCTTACCAGGCAATG	2844
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	2845	ATGTATTTTTCCTATGCTGATGTTTCCGGTTTGGAGCTACTTGGTGGCACAATAAACC	2904
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	2905	ATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTCTTTGGTGCCATGCCGCTG	2964
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerClyAlaAlaHisLeu	926
Db	2965	GGCAAGTCAGTTCATTGCTCTGACTATGCCAAGCCAAAATATCAGACGCCACCATC	3024
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysPro	946
Db	3025	ATCATGATCATTTGAAAANAACCCCTTTGATTGACGCTACAGCACCGAAGGCCCTAATCGG	3084
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966

```
Db 3085 AACACATTGGAAGAAATGTCACATTGGTGAAGTGTGATTCAACTATCCACCCGACCG 3144
Qy 967 AspValPheIleLeuArgGlyLeuSerIleLeuArgGlyLeuValAlaPhe 986
Db 3145 GACATCCAGTGTCTCAGGACCTGAGCTGGAGTGAAGAAGGCCAGACGCTGGCTCTG 3204
Qy 987 ValGlySerSerGlyCysGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db 3205 GTGGCAGCAGTGGTGTGGAGAGACACAGTGTCCAGTCTCTGGAGCGGTCTAGCAC 3264
Qy 1007 ProValGlnGlnValLeuPheAspGlyValAlaAlaLysGluLeuAsnValGlnTyr 1026
Db 3265 CCCTTGGCAGGAAAGTGTCTGTATGGTCAAAAGAAATGAAGCGACTGAATGTCAGTGG 3324
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnProValLeuPheAsnCysSerIleAla 1046
Db 3325 CTCGAGCACACCTGGGCATCGTGTCCAGAGCCCATCTGTTGACTGCAGCATGCT 3384
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db 3385 GAGAACATTCCTATGGAGAACACAGCCGGTGTGTCCACAGGAAGATCGTGAGGCCA 3444
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db 3445 GCAAAGAGGCCAACATACATGCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnArgLeuAlaIleAlaArg 1106
Db 3505 GTAGGAGACAAAGAACTCAGCTCTCTGGTGGCCAGAAACAGCGCATTCGATAGCTGT 3564
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3565 GCCTTGTAGACAGCTCATATTTTGTCTTTGTGATGAAGCCAGCTCAGCTCTGGATACA 3624
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaAlaArgThrGlyArgThrCysLeu 1146
Db 3625 GAAAGTGAAAGGTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCGCCGACCTGCATT 3684
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 3685 GTGATTGCTCAGCGCTGTCCACCATCCAAATGCAGACTTAATAGTGGTGTTCAGAAT 3744
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 3745 GCAGATCAAGAGCATGTGCACGCATCAGCTGTCTGCACAGAAAGCATCTATTTT 3804

Qy 1187 LysLeuValAsnAlaGlnSer 1193
ID ABA94366
Db 3805 TCAATGTCAGTGTCCAGGCT 3825

RESULT 12
ABA94366
AC ABA94366;
XX
XX
XX 26-MAR-2002 (first entry)
XX
DE Human BCRP DNA related seq id No. 3.
XX
XX Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatocellular; neurodegenerative; cytotostatic; antianemic; muscular; BCRP;
KW cardiac; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 1..3843
FT /*tag= a
XX
XX WO200192877-A2.
```

```
PD 06-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US17459.
XX
XX 31-MAY-2000; 2000US-0584586.
PR 29-MAY-2001; 2001US-086866.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Sorrentino B, Schuetz J;
XX
XX WPI; 2002-114368/15.
DR P-PSDB; ABB07267.
XX
XX Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, Parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell -
XX
XX Disclosure; Page 59-60; 87pp; English.
XX
XX The invention provides a method of identifying and/or isolating a stem
CC cell that involves detecting the expression of an ATP transport protein
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC comprising stem cells. The isolated stem cells may be used in the
CC treatment of diseases such as muscular dystrophy, degenerative liver
CC disorder, myocardial infarction, Parkinson's disease, degenerative liver
CC disorders of the brain, and for tissue regeneration or replacement.
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC treatment of leukemia) and for ex vivo gene therapy for treating blood
CC diseases such as sickle cell anemia and thalassemia. The stem cells can
CC also be used as cell targets in gene therapy protocols. The present
CC sequence represents a sequence related to the BCRP for which no relevant
CC information has been provided in the specification.
XX
SQ Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 3860
Score: 2545.00 Matches: 681
Percent Similarity: 75.53% Conservatives: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 24 Gaps: 9

US-09-873-409-6 (1-1195) x ABA94366 (1-3860)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTGGAGAAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAGATCTGATGTCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAGAGTGCATATCATGATACAGGGTTCTTCATGATCTGGAGGAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCAGATATGCTTATTATACAGTGGAAATGGTGGTGGTGGTGGTGGTGGTGGTGGT 390
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGTTTCAATTTTGGTGCCTGGCAGCTGGGAAGCAAAATACAAAAATACAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTAATATGACAGGAGATAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGGAG 510
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
```

Db 511 CTTAACACCGGACTTACAGATGATGCTCTCTAAGATTAAATGAAGTTATTGGTGACAAAATT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 630  
Qy 141 GlyTyrLeuLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GGTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGACGTGCT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 691 GTCGGGCAAGATACCTATCTTCATTCTATCTGATTAAGAACTCTTAGCGTATGCAAAAGCT 750  
Qy 181 GlyAlaValAlaGluLeuValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGACAGTAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATTGCAATTTGGAGGACAA 810  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysPheGlyIleLys 220  
Db 811 AAGAAAGAACTTGAAAGGTACAAACAAAATTTAGAAGAGCTAAAGAAATTTGGGATAAAG 870  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 871 AAAGCTATTACGCCAATATTCTATAGGTGCTGCTTTCTGCTGATCTATGATCTTAT 930  
Qy 241 GlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCTCTCAGGGGAA-----TATTCATT 984  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 985 GGACAGTAGTACTCACTGATTCTTCTGTATTAAATTTGGGCTTTTAGTGTGGCAGGCA 1044  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1045 TCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGGAGCAGCTTAAATCTTCAAGATA 1104  
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1105 ATTGATAAATAAGCAAGTATTGACAGTATTGCAAGAGTGGGCACAAACAGATATATT 1164  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAGAAAGTTAAG 1224  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1225 ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCTGGTGGAAAC 1284  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1285 AGTGGCTGTGGGAGAGCACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1344  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1345 GCGATGCTCAGTGTTCATGACAGAGGATATTAGGACCAATAATGTAAGGTTTCTACCGGAA 1404  
Qy 401 HisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsnAnIle 420  
Db 1405 ATCATTTGGTGTGGTAGTCAGGAACCTGATTGTTTGGCACCACGATAGCTGAAAACATT 1464  
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1465 CGCTATGGCGGTGAAATGTCCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
Db 1525 GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGACACCCCTGCTGGAGAGAGGG 1584  
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1585 GCCCAGTTGAGTGGTGGGAGAGCAGAGGATCGCCATTGCGCTGCCCTGCTTGCACAC 1644

Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1645 CCCAAGATCCTCTGCTGGATGAGGCCAGCTCAGCTTGACACAGAAAGCGAAGCAGTG 1704  
Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
Db 1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAGGTGCGACCACCATTTGTGTAGCTCATCGT 1764  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 1765 TTGCTACAGTTTGGTAATGCTGACGTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
Db 1825 AAAGGAAATCATGATCACTCATGAAGAGAAAGGCAATTTACTTCAAACTTGTCAACAATG 1884  
Qy 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
Db 1885 CAGACAGCAGGAAATGAAGTTGAATTTAGAAATGTCAGCTGATGAATCCAAAGTGAAT 1944  
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
Db 1945 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACT 2004  
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607  
Db 2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2064  
Qy 608 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 627  
Db 2065 GATGAAAGTATACCTCCAGTTTCTCTTTGGAGGATTATGAAGCTAAATTTAACTGAATG 2124  
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
Db 2125 CCTTATTTGTTGGTGTATTGTCGCATTATAAATGGAGGCTCGAACGACGATTT 2184  
Qy 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 2185 GCAATAATATTTTCAAGAGTATAGGGGTTTTTACAAAGAAATTTGATGATCTCGAAACAAA 2244  
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
Db 2245 GCACAGAAATAGTAACCTGTTTCTACTATTGTTTCTAGCCCTTTGGAATTTATTTCTTTAT 2304  
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706  
Db 2305 ACATTTTCTTCAAGGTTTCACTTTGGCAAGCTGGAGAGATCTCACCAGCGGCTC 2364  
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 726  
Db 2365 CGATACATGTTTTCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCATAA 2424  
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
Db 2425 AACACCATGGAGCATTGACTACCAAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT 2484  
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
Db 2485 ATAGGTTCCAGGCTTGGTGAATTTACCCAGAAATATAGCAAAATCTTTGGGACAGAAAT 2544  
Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786  
Db 2545 ATATCTTCATCTATGTTGGCACTACACTGTTACTTCTAGCAATTTGATCCCATCAT 2604  
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
Db 2605 GCAATAGCAGGAGGTTGTTGAAATGAAATGTTGCTGGCAAGCACTGAAAGATAAGAAA 2664  
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
Db 2665 GAACTAGAAGGTCGTGGGAGATCGCTACTGAAGCAATAGAAAACTTCCGAAACCGTTGTT 2724



```
QY 827 SerLeuThrArgGluTysAlaPheGluGlnMetTyrGluGlnMetLeuGlnThrGlnHis 846
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB TCATTGACTCAGGACGAGAAAGTTTGAACATATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB AGAAACTCTTTGAGGAAGACACATCTTTGGGAATATACATTTTCCTTCACCCAGGCAATG 2844

QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB ATGATATTTTCCATATGCTGATGTTCCGGTTTGGAGCCTACTTGTGGCACATAAATC 2904

QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB ATGAGCTTTCAGGATGTTCTGTAGTATTTTTCAGCTGTTCTGTGTGTCATGGCGGTG 2964

QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GGGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAAGCCCAAAATATATCAGCAGCCACATC 3024

QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB ATCATGATCATTTGAAAACCCCTTTGATTTGACAGCTACAGCAGCGAAGCCCTAATGCCG 3084

QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB AACACATTCGAAGGAATGTACATTTTGGTGAAGTTGATTTCAACTATCCACCCGACCG 3144

QY 967 AspValPheIleLeuArgGlyLeuSerIleLeuArgGlyLysThrValAlaPhe 986
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GACATCCAGCTGCTTTCAGGACTGAGCTCGAGGTGAAGAAGGCCAGCGCTGGCTCTG 3204

QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuArgLeuTyrAsp 1006
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GTGGCAGCAGTGGCTGTGGAGAGACACAGTGTCTCAGCTCTCTGGAGCGGTCTTACGAC 3264

QY 1007 ProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB CCCTTGGCAGGAAAGTGCTGTGTATGTGTCGAAAGAAATAAAGCGACTGAATGTTCACTGG 3324

QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB CTCGAGACACCTGGGCATCGTGTCCAGAGGCCATCTCTGTGACTCGACATGCT 3384

QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GAGNACATTGCTATGAGACACACAGCCGGTGTGTACAGAGAGATCTGTAGGGCA 3444

QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GCAAGGAGGCCAACATACATGCTTTCATCGAGTCACTGCTTAATAATATAGCACTAA 3504

QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GTAGGAGACAAAGAACTCAGCTCTCTGTGGTGGCCAGAAACAGCCATTCGATGCTCGT 3564

QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GCCCTGTTAGACAGCTCATATTTTGGATGAGCCAGCTGCTGATGATACA 3624

QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3684

QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuValValLeuHisAsn 1166
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GTGATTCCTCAGCGCTGTCACCATCCAGATGCAGACTTAATAGTGGTGTTCAGAAT 3744

QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGlnLeuArgAsnArgAspIleTyrPhe 1186
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB GGCAGATCAAGAGCATGTCAGCATCAGCAGCTGTGTCACAGAAAGGCATCTATTTT 3804

QY 1187 LysLeuValAsnAlaGlnSer 1193
```

```
DB 3805 TCAATGTCAGTGTCCAGGCT 3825
RESULT 13
AAH57442
ID AAH57442 standard; cDNA; 4349 BP.
XX AAH57442;
XX 10-SEP-2001 (first entry)
XX Human intestine cell specific cDNA sequence SEQ ID NO:282.
XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
XX lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
XX metabolic disease; developmental disease; cytostatic; immunomodulatory;
XX neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX Homo sapiens.
XX WO200132927-A2.
XX 10-MAY-2001.
XX 02-NOV-2000; 2000WO-US30396.
XX 04-NOV-1999; 99US-0163508.
XX (INCY-) INCYTE GENOMICS INC.
XX Sornasse T, Seilhamer JJ, Watson GA;
XX WPI; 2001-291057/30.
XX New cell and tissue specific polynucleotides useful for diagnosis, is
XX prognosis or monitoring of treatments for disorders where the gene is
XX associated with a cancer, immunopathology or neuropathology -
XX Claim 1; Page 207-208; 327pp; English.
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
XX sequences (I). (I) can have cytostatic, immunomodulatory and
XX neuroprotective activities, and can be used in gene therapy. (I) and
XX proteins (II) encoded by then are used in high throughput screening
XX assays to select DNA molecules, RNA molecules, peptide nucleic acids,
XX mimetics, peptides, proteins, agonists, antagonists, antibodies or
XX their fragments, immunoglobulins, inhibitors, drug compounds and
XX pharmaceutical agents. Expression of (I) in a sample indicates the
XX differentiation of embryonic stem cells into a tissue selected from
XX brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
XX tissues. (I) and (II) are used to produce an expression profile that
XX defines a metabolic or developmental process, treatment, condition,
XX disease or disorder. The gene profile can be used for diagnosis,
XX prognosis or monitoring of treatments and for investigating a
XX predisposition to a disorder where the gene is associated with a
XX cancer, immunopathology or neuropathology.
XX Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;
Alignment Scores:
Pred. No.: 0 Length: 4349
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 22 Gaps: 9
US-09-873-409-6 (1-1195) x AAH57442 (1-4349)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 276 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCTGAGGCTGGAGCTTCTCTCATGATGCTG 335
```



Qy	21	ValLeuGlyGluMetSerAsp	-----AsnLeuIleSer	31
Db	336	GTGTTGGAGAAATGCAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTC	395	
Qy	32	GlyCysLeuValGlnThr	-----AsnThrTyrSerPhePhe-	43
Db	396	AACATCACTAATAGAGTGTATCAATGATACAGGGTCTCTCATGAATCTCGAGGAGAC	455	
Qy	44	-----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr	61	
Db	456	ATGACCAAGGTATGCTATTATTACAGTGGAAATGGTCTGGGGTCTGGTGGCTCTAC	515	
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81	
Db	516	ATTTCAGGTTCATTTTGGTGCCTCGGAGCTGGGAACAATAACACAAATTTAGAAAAACAG	575	
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyClu	101	
Db	576	TTTTTTTCATGCTAATATGCGACAGGAGTAGGCTGGTTTGATGTCACAGTGTGGGGAG	635	
Qy	102	LeuAsnThrArgMetThr	---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	636	CTTAAACACCCGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTGGTGACAAAATT	695	
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys	140	
Db	696	GGAAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATATAGGATTTACACGT	755	
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160	
Db	756	GGTTGGAAAGCTAACCCCTTGATTTTGGCCATCAGTCTGCTTTCTTGGACTGTGAGTGTCT	815	
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180	
Db	816	GTCTGGGCAAGATACATCTTCTTCAATTTCTGATAAAGAACTCTTACGCTATGCAAAAGCT	875	
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200	
Db	876	GGAGCAGTAGCTGAAGAGTCTTTGGCGAAATTAGAACTGTGATTGCAATTTGGAGGACAA	935	
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220	
Db	936	AAGAAAGAACTTGAAGAGGTACAAACAAATTTAGAAAGAGCTTAAAGAAATTTGGGATAAAG	995	
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyrTyr	240	
Db	996	AAAGCTATTACAGCAATATTCTATAGGTGCTGCTTTCTGCTCATCTATGCATCTTAT	1055	
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260	
Db	1056	GCTCTGGCCTTCTGGTATGGACCACTTTGGTCTCTCAGGGGAA-----TATTCTATT	1109	
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280	
Db	1110	GGACAAGTACTCACTGATTCTTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA	1169	
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300	
Db	1170	TCTCCAGCATTTGACAAATTCGACAGGAGCGAGCTTATGAATCTTCAAGATA	1229	
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320	
Db	1230	ATTGATTAATAAGCCAAGTATTGACAGCTATTCCAGAGAGTGGGCACAAACCAGATAATATT	1289	
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340	
Db	1290	AAGGGAAATTTGGAATTCGAAATGTTCTACATTTCAGTTACCCATCTCGAAAAAGAGTTAAG	1349	
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360	
Db	1350	ATCTTGAAGGCCTGAACCTGAAGGTGCGAGGTGGCGACGGTGGCCCTGGTTGGAAAC	1409	
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnAtgLeuTyrAspProAspAsp	380	

[illegible]

Db 2490 CGATACATGGTTTCCGATCCATGCTCAGACAGGATGTGATGGTTGGTTGATGACCCCTAAA 2549  
QY 727 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyVala 746  
Db 2550 AACACCACTGGAGCATTTGATACAGGCTCGCCAAATGATGCTCAAGTTAAAGGGCT 2609  
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValile 766  
Db 2610 ATAGGTTCCAGGCTGTGTAAATTACCCAGAAATATAGCAAAATCTTGGGACAGGAATAATT 2669  
QY 767 IleSerPheIleThrGlyTyrGluMetThrPheIleLeuSerIleAlaProValLeu 786  
Db 2670 ATATCTTCATCTATGTTGGCAATACCACTGTATCTTACCAATTTACCCATCAT 2729  
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
Db 2730 GCAATACAGAGAGTTGTGAATGAATAATGTTGTCTGGACAGCACTGAAGAATAGAAA 2789  
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
Db 2790 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAAACCGTTGTT 2849  
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846  
Db 2850 TCTTTGATCAGGAGCAGAAAGTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 2909  
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866  
Db 2910 AGAACTCTTTAGGAAGACACACATCTTTGGAATACATTTCCCTTCACCCAGGCAATG 2969  
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleLeuGlnAlaGlyArg 886  
Db 2970 ATGTATATTTTCTATGCTGGATGTTCCGGTTTGGAGCTTACTTGGTGGCAATAAACTC 3029  
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
Db 3030 ATAGCTTTGAGGATGTCCTGTAGTATTTTCAGCTGTTGCTGTGGTGCATGGCGGTG 3089  
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926  
Db 3090 GGGCAAGTCAGTTCAATTTGCTCTGACTATGCCAAGCCAAATATCAGAGCCCATC 3149  
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946  
Db 3150 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAAAGCCCTAATGCCG 3209  
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
Db 3210 AACACATTGGAAGGAATGTCAATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3269  
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986  
Db 3270 GATATCCAGTGCTTCAGGACTTGAGCTGGAGGTGAGAGGGCCAGACGCTGGCTCTG 3329  
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006  
Db 3330 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCTCCAGCTCTGGAGCGGTTCTACGAC 3389  
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026  
Db 3390 CCTTGGCAGGAAAGTGTCTGTATGGCAAGAAATAAAGCGACTGAATGTTCAGTGG 3449  
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046  
Db 3450 CTCGAGCACACCTGGGCATCTGTCTCCAGAGGCCCATCTCTGTTGACTGCAGCATGCT 3509  
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066  
Db 3510 GAGAACTTCCCTATGAGACAAACAGCCGGTGTGTGCACAGGAAGAGATCGTGGAGGCA 3569  
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086  
Db 3570 GCAAGAGGAGCCCAACATACATGCTTTCATGATCTGCTTAAATAAATAGCACTAAA 3629

QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106  
Db 3630 GTAGGAGACAAAGAACTCAGCTCTCTGTGGCCAGAAACAGCATTGCCATAGCTCGT 3689  
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 3690 GCCCTTTGTAGACAGCTCATATTTTGTCTTTGGATGAAGCCACGTCAGCTCGGATACA 3749  
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
Db 3750 GAAAGTGAAAGGTTTCCAAAGAACCCCTGGCAAGCCAGAGAGCCGACCTGCATT 3809  
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166  
Db 3810 GTGATTGCTCACCGCTCTCCACCATCCAGATGAGACTTAATAGTGTGTTTCAGAAAT 3869  
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186  
Db 3870 GGCAGAGTCAAGGAGCATGGCAGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3929  
QY 1187 LysLeuValAsnAlaGlnSer 1193  
Db 3930 TCAATGTCAGTGTCCAGGCT 3950  
RESULT 14  
AAZ94738  
ID AAZ94738 standard; cDNA; 4646 BP.  
XX  
AC AAZ94738;  
XX  
DT 01-AUG-2000 (first entry)  
XX  
DE Human ATP binding cassette ABCB1 (MDR1) cDNA.  
XX  
KW ABCB1; ATP binding cassette; human; cholesterol; lipid disorder; atherosclerosis; lipid disorder; dyslipidemia; psoriasis;  
KW lupus erythematosus; diagnosis; gene therapy; MDR1;  
KW multidrug resistance; chromosome 7q21, ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200018912-A2.  
XX  
PD 06-APR-2000.  
PF 21-SEP-1999; 99WO-EP06991.  
XX  
PR 25-SEP-1998; 98US-0101706.  
XX (FARB ) BAYER AG.  
PA Schmitz G, Klucken J;  
PI  
XX  
DR WPI; 2000-293151/25.  
PT Adenosine triphosphate binding proteins useful for identifying agents for treating atherosclerosis and other inflammatory disorders -  
XX  
PS Claim 9; Page 110-112; 154pp; English.  
CC The present sequence is that of human ATP binding cassette subfamily B protein ABCB1 cDNA. The cDNA was identified using a differential display method in which monocytes from peripheral blood were subjected to macrophage differentiation and cholesterol loading with acetylated low density lipoproteins and subsequent deloading with high density lipoprotein (HDL3) to identify cholesterol sensitive genes. The gene maps to chromosome 7q21 and is also termed MDR1 (multidrug resistance). The invention provides cholesterol-sensitive ABC genes (see AAZ94734-63). These genes, and polypeptides encoded by them, can be used for diagnostic and therapeutic applications, and for biochemical or cell-based assays to screen for pharmacologically active modulator compounds

CC useful for the treatment of lipid disorders, atherosclerosis or  
CC other inflammatory diseases such as psoriasis and lupus  
XX erythematous.

SO Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

## Alignment Scores:

Pred. No.:	0	Length:	4646
Score:	3545.00	Matches:	681
Percent Similarity:	75.55%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	21	Gaps:	9

US-09-873-409-6 (1-1195) x AAZ94738 (1-4646)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	575	ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCTCATGATGCTG	634
Qy	21	ValLeuGlyGluMetSerAsp-----AsnLeuIleSer	31
Db	635	GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAAAGATCTGATGTCA	694
Qy	32	GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-----	43
Db	695	AACATCATCTAATAGAAAGATATCATATGATATACAGGGTTCTTCATGAATCTGGAGGAAGAC	754
Qy	44	-----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr	61
Db	755	ATGACAGGATATGCTTATTATACAGTGAATTTGGTCTGGGGTGTCTGTTCTGCTTAC	814
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
Db	815	ATTCAAGTTTCACTTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATATTAGAAACAG	874
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu	101
Db	875	TTTTTTCATGCTATTAATCGCAGCAGAGATAGGCTGGTTTGTATGTGCACCATGTGGGGAG	934
Qy	102	LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	935	CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGTGCACAAAT	994
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	995	CGAATGCTCTTTTCAGTCATGCAACATTTTTCACCTGGGTATTATAGTAGATTACAGCT	1054
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	1055	GGTTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCACTGCT	1114
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	1115	GTCTGGGCAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATCAAAAGCT	1174
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	1175	GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGACTGTGATGCAITTTGGAGACAA	1234
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	1235	AAGAAGAACTTTGAAAGGTACAAACAAATTTTAGAAGAGCTAAAGAAATTTGGGATAAG	1294
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	1295	AAAGCTATTACGCCAATATTCTTATAGTGTGCTTCTCTGCTGATCATCATCTTAT	1354
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	1355	GCTCTGGCCCTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT	1408
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280

Db	1409	GGCAAGTACTCACCTGTATTTCTGTATTAATTTGGGGCTTTTAGTGTGGACAGGCA	1468
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	1469	TCCTCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA	1528
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	1529	ATTGATAATAAGCCAAAGTATTGACAGCTATTTCGAAGAGTGGGCAAAACAGATAATAT	1588
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	1589	AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG	1648
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1649	ATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGACGCTGCTGTTGGAAC	1708
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	1709	AGTGGCTGTGGAGAGCACAACAGTCCAGCTGTATGCAGAGGCTCTATGACCCACAGAG	1768
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	1769	GGGATGGTCAGTGTGTGATGGACAGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA	1828
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	1829	ATCATGTGTGTGTGAGTCAGGAACTGTATTGTTTGGCCACCATAGTAGTCAAAACATT	1888
Qy	421	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1889	CGCTATGCCGCGTAAATGTCCACATGATGATGTAGAGAGCTGTGAGAAAGCAAT	1948
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1949	GGCTATGACTTTTATCATGAACCTGCTCATAAATTTGACACCTGTTGGAGAGAGAGGG	2008
Qy	461	AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	2009	GCCCAGTTGAGTGGTGGCAGAGCAGAGGATCGCCATTCGACGTGCTGTTGCCAAC	2068
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	2069	CCCAAGATCCTCCTGCTGGATGAGCCACGTCAGCCTTGGACACAGAAAGCAAGCAGTG	2128
Qy	501	ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	2129	GTTCAGGTGGCTCTGGTAAGGCCAGAAAAGGTCGGACCCACCATTTGTATAGCTCATCGT	2188
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	2189	TGTCTACAGTTCTGTAATGCTGACGTCACTCGCTGCTGTTTCGATGATGGAGTCAITTTGGAG	2248
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
Db	2249	AAAGGAATCATGATGAATCATGAAAGAGAAAGGCAITTTACTTCAAACTTGTCACAATG	2308
Qy	561	Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr	574
Db	2309	CAGACAGCGAAATGAAGTTGAATTAGAAAATGTCAGCTGATGAATCCAAAAGTGAAT	2368
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	2369	GATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGATCCAGTCTTAATAAGAAAAGATCACT	2428
Qy	590	---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys	607
Db	2429	CGTAGGAGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAGAGGCTCTG	2488
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp	627

Db 2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTTAAAGCTAAATTTAACTGAATGC 2548  
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
Db 2549 CCTATTATTTGCTGGTGTATTTGTGCATTTATTAATGGAGGCTGCAACAGCATTT 2608  
Qy 648 SerIlePheAlaLysIleLeuThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 2609 GCATATATATTTCAAGATTTATAGGGTTTTTACAGAAATGATGATCCTGAAACAAA 2668  
Qy 667 LysHisAspAlaGluLeuLysSerMetIlePheValLeuGlyValIleCysPheVal 686  
Db 2669 CGACAGAAATAGTAATGTTTCTACTATTGTTCTAGCCCTTGAATATATTTCTTTAT 2728  
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyValLeuThrMetArgLeu 706  
Db 2729 ACATTTTCTCCAGGGTTTCACATTTGGCAAGCTGGAGAGATCTCCAAAGCGGCTC 2788  
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaThrPheAspGluLysGlu 726  
Db 2789 CGATACATGTTTCCGATCCATGCTCAGACAGATGTGAGTTGGTTGATGACCTAAA 2848  
Qy 727 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
Db 2849 AACACCACTGGAGCATGACTACAGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT 2908  
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
Db 2909 ATAGGTTCCAGGCTGCTGTAATTTACCAAGATATAGCAATCTTGGGACAGGAATAAT 2968  
Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786  
Db 2969 ATATCTTCTATCTATGTTGGCACTTAACACTGTACTCTTAGCAATGTACCCATCAT 3028  
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
Db 3029 GCAATAGCAGGAGTTGTTGAAATGAAATGTTGCTGGCAAGCACTGAAAGATAAGAAA 3088  
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
Db 3089 GAACATAGAGGTCGTGGGAAGATCGCTACTGGAAGCAATAGAAACTTCCGAACCGTTGT 3148  
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 846  
Db 3149 TCITTGACTAGGAGCAGAGATTGAAATATGATGCTCAGAGTTTGCAGGTACCATAC 3208  
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866  
Db 3209 AGAACTCTTTGAGGAAAGCACACATCTTTGGAATTAATTTTCTTCCATCCAGGCAATG 3268  
Qy 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886  
Db 3269 ATGTAATTTTCTATGCTGATGATGTTCCGGTTTGGAGCCTACTTGTGGCACATAAATC 3328  
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
Db 3329 ATGAGCTTGAGGATGTTCTGTAGTATTTTACACTGTTCTTGTGTGCATGGCCGTG 3388  
Qy 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926  
Db 3389 GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAATATCAGCAGCCACATC 3448  
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946  
Db 3449 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCAGGAAGGCTTAATCCG 3508  
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
Db 3509 AACACATTTGGAGGAATGTGCATTTGGTGAAGTTGATTTCAACTATCCACCCGACCG 3568  
Qy 967 AspValPheIleLeuArgGlyLeuSerIleGluArgGlyLysThrValAlaPhe 986  
Db 3569 GACATCCAGTGCTTCAGGGACTGAGGCTGGAGGTGAAGAGGCGCAGACGCTGGCTCTG 3628

Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006  
Db 3629 GTGGGACAGAGTGGCTGTGGGAAGACAGCAGTGGTCCAGCTCTCTGGAGCGGTTCACGAC 3688  
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026  
Db 3689 CCTTGGCAGGAGAAAGTGTCTGTATGTCGCAAAAGAAATAAAGCGACTGAATGTTCAAGTG 3748  
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046  
Db 3749 CTCGAGACACACTGGGCATCTGTGCCAGAGGCCCATCTCTGTTCAGCTGCAGCATGCT 3808  
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066  
Db 3809 GAGACATTTCCCTATGGACACAACAGCCGGTGGTGTCCACAGGAAGAGATCGTGAGGGCA 3868  
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086  
Db 3869 GCAAGGAGGCGCAACATACATGCTTTCATCGAGTCACCTGCTTAATAATAATATAGCACTAA 3928  
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 1106  
Db 3929 GTAGGAGACAAAGGAACCTCAGCTCTCTGTGGCCAGAAACAGCATTCCTAGTCTCGT 3988  
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 3989 GCCCTGTGTAGACAGCTCATATTTGCTTTTGGATGAAGCCAGCTCAGCTCTGGATACA 4048  
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
Db 4049 GAAAGTGAAAGGTTTGTCCAAGAAAGCCCTGGCAAAAGCCAGAGAGCGGCACCTGCAT 4108  
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166  
Db 4109 GTGATTCCTCAGCGCTGTCACCATCCAGATTCAGACTTAATAGTGGTGTTCAGAAAT 4168  
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186  
Db 4169 GGCAGAGTCAGAGAGATGGCAGCATCAGAGCTCTGTCACAGAAAGGCATCTATTT 4228  
Qy 1187 LysLeuValAsnAlaGlnSer 1193  
Db 4229 TCAATGTCAGTGTCCAGGCT 4249  
RESULT 15  
AAD38994  
ID AAD38994 standard; DNA; 4646 BP.  
XX  
AC AAD38994;  
XX  
DT 23-SEP-2002 (first entry)  
XX  
DE Human mdrl gene.  
XX  
KW Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;  
KW hypoxia inducible factor-1; small ubiquitin-like-modifier; HIP-1;  
KW lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;  
KW myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;  
KW angiotenic myeloid metaplasia; myeloid leukaemia; gene therapy;  
KW polycythaemia vera; hypoxia responsive element; HRE; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 425..4267  
FT /tag= a  
FT /product= "Human MDR protein"  
XX  
PN WC020234291-A2.  
XX  
PD 02-MAY-2002.  
XX

PF 25-OCT-2001; 2001WO-US49856.

PR 26-OCT-2000; 2000US-243542P.

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.

XX Colgan SP;

XX WPI; 2002-471427/50.

XX P-PSDB; AAE24211.

DR Treating a subject (at risk of) having a hematologic malignancy or

XX multidrug resistance, e.g. lymphoma or myeloma, by administering

PT hypoxia inducible factor 1 binding molecules or small

PT ubiquitin-like-modifier-1 binding molecules -

XX Disclosure; Page 53-55; 92pp; English.

PS The invention relates to a method of treating a subject having or at

XX risk of developing a haematologic malignancy or multidrug resistance

CC (MDR). The method involves administering hypoxia inducible factor-1

CC (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1

CC binding molecules or HIF-1-SUMO-1 complex modulators. mdrl-hypoxia

CC responsive element (HRE) binding molecules or antisense nucleic

CC acid molecules and SUMO-1 binding molecules or antisense molecules

CC are useful for treating a subject having or at risk of developing

CC haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid

CC disorder). The lymphoid disorders include lymphocytic leukaemia or

CC chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic

CC lymphoid leukaemia. The myeloid disorders include chronic or acute

CC myeloid leukaemia, e.g. angioleukemia myeloid metaplasia, essential

CC thrombocythaemia or polycythaemia vera. The invention is used in gene

CC therapy. The present sequence is human mdrl gene.

XX SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

#### Alignment Scores:

Pred. No.:	0	Length:	4646
Score:	3545.00	Matches:	681
Percent Similarity:	75.55%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	24	Gaps:	9

US-09-873-409-6 (1-1195) x AAD38994 (1-4646)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 575 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATCGGGCTGGACCTTCTCATGATGCTG 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 635 GTGTTGGAGAAATGACAGATATCTTGGCAATGACAGGAATTTAGAAATCTGATGTC 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43

Db 695 AACATCACTAATAGAAAGTATGATCAATGATACAGGCTTCTTCAATGATCTGGAGGAAC 754

Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 51

Db 755 ATGACACCATGATGCTTATTATACAGTGAATTTGGTGGGCTGGTGGTGGTGGTGGTGG 814

Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgIleGlyGln 81

Db 815 ATTCAGTTTTCATTTTGGTGGCTGGCAGCTGGAGAACAAATACACAAATTTAGAAACAG 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101

Db 875 TTTTTCATGCTATATGCGACAGGAGATAGGCTGTTTGTATGTCGACCATGTTGGGGAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120

Db 935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys 140

Db 995 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCATCGGGTTTATAGTAGGATTTACACGT 1054

Qy 141 GlyTrpIysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160

Db 1055 GGTTCGGAAGCTAACCCCTTGTGATTTTGGCCATCACTCTCTTCTTGGAGCTGTCAGCTGCT 1114

Qy 161 AlaCysSerArgMetValIleSerLeuThrSerIysGluLeuSerAlaTyrSerIysAla 180

Db 1115 GTCCTGGGCAAGATACTATCTTCACTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

Db 1175 GCGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGAACTGTGATTCATTGGAGACAA 1234

Qy 201 GluIysGluLeuGlnArgTyrThrGlnAsnLeuIysAspAlaIysAspPheGlyIleIys 220

Db 1235 AAGAAGAACTTTGAAAGGTACAACAAAATTTAGAAAGCTTAAAGAAATTTGGGATAAAG 1294

Qy 221 ArgThrIleAlaSerIysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240

Db 1295 AAAGCTATTACAGCAATATTTCTATAGTGTCTTCTCTCTGCTGATCTATCATCTTAT 1354

Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260

Db 1355 GCTCTGGCTCTTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1408

Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280

Db 1409 GCGAAGTACTCACTGTATCTTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1468

Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300

Db 1469 TCTCAAGCATTTGAAGCATTTGCCAATTCAGAGGAGAGCTTATGAAATCTTCAGATA 1528

Qy 301 IleAspIysLysProSerIleAspAsnPheSerThrAlaGlyTyrIysProGluSerIle 320

Db 1529 ATTGATAATAAGCAAGTATTTCAGAGCTATTTCGAAGAGTGGGCAACAAACAGATAATAT 1588

Qy 321 GluGlyThrValGluPheIysAsnValSerPheAsnTyrProSerArgProSerIleIys 340

Db 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAAGTTAAG 1648

Qy 341 IleLeuIysGlyLeuAsnLeuArgIleIysSerGlyGluThrValAlaLeuValGlyLeu 360

Db 1649 ATCTTGAAGGGCTGAACTGAAAGTGCAGAGTGGGACAGCGTGGCTGTTGGAAAC 1708

Qy 361 AsnGlySerGlyIysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380

Db 1709 AGTGGCTGTGGGAAGAGCACACAGCTCCAGCTGATGTCAGAGGCTCTATGACCCACAGAG 1768

Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400

Db 1769 GGGATGGTGCAGTGTTCATGGCAGGAGATATTAGGACCATATAAATGTAAGGTTTCTACGGGAA 1828

Qy 401 HisIleGlyValIysSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420

Db 1829 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1888

Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440

Db 1889 CGCTATGGCGTGAATGTCACCATGGATGATGATGATGATGATGATGATGATGATGATGATG 1948

Qy 441 AlaTyrAspPheIleMetGluPheProAsnIysPheAsnThrLeuValGlyGlyIysGly 460

Db 1949 GCCTATGACTTTTATCATGAACTGCTCATATAATTTGACACCTGCTGTTGGAGAGAGGG 2008

Qy 461 AlaGlnMetSerGlyGlnIysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480

Db 2009 GCCCAGTTGAGTGGTGGGAGAGAGGATCGCCATTTGACGCTGCCCTGGTTCGCAAC 2068



Tue Apr 1 09:38:09 2003

us-09-873-409-6.p2n.rng

Page 35

Db 4229 TCNATGTCAGTGTCCAGGCT 4249

Search completed: March 30, 2003, 03:20:36  
Job time : 1405.91 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 8219.42 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-6  
Perfect score: 6012  
Sequence: 1 MILGILASLVNGACPLPLMLP.....QELLNRDIYFKLVNAQSVQ 1195

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-MODEL=frame\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US09873409/runat\_27032003\_115420\_19240/app\_query.fasta\_1.7544  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1.1 30544 @runat\_27032003\_115420\_19240 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estov:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_huv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_ocher:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	18.2	669	13	BM013981	BM013981 603639619
2	1051	17.5	1033	14	BM926413	BM926413 AGENCOURT
3	1008	16.8	943	9	AL520322	AL520322 AL520322
4	995	16.6	760	13	BM015204	BM015204 603642659
5	926.5	15.4	929	14	BQ882401	BQ882401 AGENCOURT
6	918	15.3	545	9	AL040762	AL040762 DKF2P434C
7	856.5	14.2	948	13	BM471690	BM471690 AGENCOURT
8	854	14.2	2676	11	AK014319	AK014319 Mus muscu
9	834	13.9	998	14	BM904842	BM904842 AGENCOURT
10	771	12.8	894	12	BF584668	BF584668 602098406
11	770.5	12.8	1019	12	BG248052	BG248052 602359987
12	753	12.5	726	12	BG293345	BG293345 602390738
13	752.5	12.5	871	17	AZ682350	AZ682350 ENTKB16TF
14	751	12.5	944	12	BF796582	BF796582 602258463
15	744.5	12.4	872	11	AK020318	AK020318 Mus muscu
16	743.5	12.4	886	17	BH139685	BH139685 ENTNA47TR
c	734.5	12.2	932	17	AZ670821	AZ670821 ENTUN69TF
17	728.5	12.1	947	17	AZ683753	AZ683753 ENTIL96TF
18	726.5	12.1	886	17	AZ540627	AZ540627 ENTKB16TF
c	723.5	12.0	880	17	AZ687805	AZ687805 ENTKB16TF
20	723.5	12.0	939	14	BQ720763	BQ720763 AGENCOURT
21	721.5	12.0	897	17	AZ541090	AZ541090 ENTDS67TR
22	720.5	12.0	897	17	BG587938	BG587938 EST489713
23	719	12.0	780	12	BH155700	BH155700 ENTKB16TF
24	711.5	11.8	913	17	AZ548312	AZ548312 ENTKB16TF
25	709.5	11.8	834	17	AZ548312	AZ548312 ENTKB16TF
c	708.5	11.8	853	17	AZ679807	AZ679807 ENTKB16TF
26	707.5	11.7	823	17	AZ532602	AZ532602 ENTKB16TF
27	702.5	11.7	1813	11	U66688	U66688 Homo sapien
28	697.5	11.6	861	14	BQ717101	BQ717101 AGENCOURT
29	697.5	11.5	1341	11	AY108285	AY108285 Zea mays
30	692.5	11.5	899	17	BH154857	BH154857 ENTKB16TF
31	690.5	11.5	835	13	BJ349604	BJ349604 B3349604
c	690.5	11.5	891	17	AZ682250	AZ682250 ENTKB16TF
33	689.5	11.5	891	17	AZ682250	AZ682250 ENTKB16TF
34	688	11.4	1123	11	AY108485	AY108485 Zea mays
35	687	11.4	715	14	BQ869512	BQ869512 QGD6G09.Y
c	684	11.4	820	13	BJ356540	BJ356540 B3356540
36	683	11.4	600	13	BM486593	BM486593 pgm2n.pk0
37	682.5	11.4	919	17	AZ690701	AZ690701 ENTKB16TF
38	680	11.3	926	14	BQ123477	BQ123477 EST609053
39	680	11.3	795	12	BF133560	BF133560 601900492
40	677.5	11.2	815	13	BI409057	BI409057 602961127
41	674.5	11.2	2564	11	AK020022	AK020022 Mus muscu
42	674	11.2	803	12	BG298756	BG298756 602396681
43	671.5	11.2	852	17	BH720383	BH720383 BOH22947F
44	670	11.1	547	13	BQ030711	BQ030711 BQ030711
45	669	11.1	547	13	BQ030711	BQ030711 BQ030711

ALIGNMENTS

RESULT 1  
BM013981  
LOCUS 603639619F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5415556 5',  
DEFINITION mRNA sequence.  
ACCESSION BM013981  
VERSION BM013981.1 GI:16528335  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 669)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTp  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12060 Row: f Column: 05  
High quality sequence stop: 664.  
Location/Qualifiers  
1. 669  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5415556"  
/clone\_lib="NIH\_MGC\_87"  
/tissue\_types="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH MGC Library."  
203 a 122 c 162 g 182 t

FEATURES  
source

Alignment Scores:  
Pred. No.: 6.69e-109 Length: 669  
Score: 1092.00 Matches: 215  
Percent Similarity: 99.54% Conservative: 2  
Best Local Similarity: 98.62% Mismatches: 0  
Query Match: 18.16% Indels: 1  
DB: 13 Gaps: 0

US-09-873-409-6 (1-1195) x BM013981 (1-669)

QY 265 AlaValPheSerValIleHisSerTyrCysIleGlyAlaAlaValProHisPhe 284  
DB 17 AGTGTTCCTTAAATGTAATCCATAGACAGTATTGCAATGGAGCAGGTCCTCAGCTTT 76  
QY 285 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 304  
DB 77 GAAACCTTCGCAATAGCCGAGGAGCTGCCTTCATATTTTCCAGGTATTGATAAGAAA 136  
QY 305 ProSerIleAspAsnPheSerThrAlaGlyTyrIlyProGluSerIleGluGlyThrVal 324  
DB 137 CCCAGTAGATAACTTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGTG 196  
QY 325 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 344  
DB 197 GAATTTAAAATGTTTCTTCAATATTCATCCATCAGACATCTATCAGATTCGAAAGGT 256  
QY 345 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 364  
DB 257 CTGAATCTCAGAAATAGTCTGGAGACAGCTGCCTTGGTGGTCTCAATGGCAGTGGG 316  
QY 365 LysSerThrValValGlnLeuLeuGlnArgIleuTyrAspProAspAspGlyPheIleMet 384  
DB 317 AAGAGTAGCGGTAGTCCAGCTCTCTCAGAGGTATATGATCCGGATGATGGCTTTATCATG 376  
QY 385 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 404  
DB 377 GTGATGAGAAATGATCATCAGAGCTTTAAATGTGGCATTTATCAGACCATATTGGAGTG 436  
QY 405 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAniIleLysTyrGlyArg 424  
DB 437 GTTAGTCAAGAGCCGTGTTTGTTCGGACCAACCACATCATGTAACAATATCAAGTATGGACGA 496  
QY 425 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 444  
DB 497 GATGATGTGATGATGAAGATGGAGAGAGCAGCAAGGGGAAGCAAAATGCCGTATGATTTT 556

QY 445 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 464  
DB 557 ATCATGAGGTTTCCTAATAAATTTAATACATTGTTAGGGGAAAAAGAGCTCAATGAGT 616  
QY 465 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 481  
DB 617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTGCTGCTTAGTTTCGAACCCC 668

## RESULT 2

BM926413 1033 bp mRNA linear EST 12-MAR-2002  
LOCUS AGENCOUT\_6600788 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5764845  
DEFINITION 5', mRNA sequence.

ACCESSION BM926413  
VERSION BM926413.1 GI:19376792  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1033)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12818 Row: o Column: 22

High quality sequence stop: 3

High quality sequence stop: 686.

## FEATURES

source

Location/Qualifiers  
1. 1033  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5764845"  
/clone\_lib="NIH\_MGC\_114"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
male brains, age range 23-27 yo. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.5 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 019. Note:  
this is a NIH MGC Library."  
289 a 212 c 218 g 312 t 2 others

## BASE COUNT

289 a 212 c 218 g 312 t

## ORIGIN

Alignment Scores:  
Pred. No.: 4.68e-104 Length: 1033  
Score: 1051.00 Matches: 221  
Percent Similarity: 94.49% Conservative: 2  
Best Local Similarity: 93.64% Mismatches: 13  
Query Match: 17.48% Indels: 3  
DB: 14 Gaps: 0

US-09-873-409-6 (1-1195) x BM926413 (1-1033)

QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
DB 180 GTTTCCTTTAGTGTAAATCCATAGCAGTTATTGATTGGAGCAGCTCCCTCAGTTGAA 239  
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 305  
DB 240 ACCTTCGCAATAGCCGAGGAGCTGCCTTTTCATATTTTCCAGGTTCATTGATAAGAAACC 299



```

|||||
Db 785 ACCATCCAGAAATGCAGACTTAAATAGTGTGTTTTCAGAAATGGCAGAGTCAAGGACATGG 844
|||||
Qy 1173 yThrHisGlnLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 1193
|||||
Db 845 CACGCATCAGAGCTGCTGGCAGAAAGGATCTATTTTCAATGGTGTCCAGGC 904
|||||
Qy 1193 r 1193
Db 905 T 905

RESULT 4
BM016204 760 bp mRNA linear EST 30-OCT-2001
LOCUS 603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
DEFINITION mRNA sequence.
ACCESSION BM016204
VERSION BM016204.1 GI:16530558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12068 row: e column: 16
High quality sequence stop: 740.
FEATURES
source
1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH MGC 87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3..76e-98 Length: 760
Score: 995.00 Matches: 203
Percent Similarity: 98.07% Conservative: 0
Best Local Similarity: 98.07% Mismatches: 2
Query Match: 16.55% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-6 (1-1195) x BM016204 (1-760)

Qy 303 LysLysProSerIleAspAsnPhSerThrAlaGlyTyrLysProGluSerIleGluGly 322
|||||
Db 3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAATCCATGATGAGGA 62
|||||
Qy 323 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 342
|||||
Db 63 ACTGTGAATTTAAATAGTTCTTTCAATATATCCATCAAGACCATCTATCAAGATTCG 122
|||||
Qy 343 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 362
|||||

```

```

|||||
Db 123 AAAGGCTGAATCTCAGAATTAGTCTGGAGAGACAGTGCCTTGGTTCGCTTCAATGGC 182
|||||
Qy 363 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPhe 382
|||||
Db 183 AGTGGGAAGAGTAGGTAGTCCAGCTCTTGCGAGAGGTTATATGATCCGATGATGGCTTT 242
|||||
Qy 383 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 402
|||||
Db 243 ATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGCATTTATCGAGACCATATN 302
|||||
Qy 403 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerPheAsnIleLysTyr 422
|||||
Db 303 GGAGTGTAGTCAAGAGCGCTGTTTGTTCGGACCAACCATCATAGTAACAATATCAAGTAT 362
|||||
Qy 423 GlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 442
|||||
Db 363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGAAGCAAAATGCGGTAT 422
|||||
Qy 443 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 462
|||||
Db 423 GATTTTATCATGAGTTTCTTAATAATTAAATACATTGGTAGGGGAAAAGAGCTCAA 482
|||||
Qy 463 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 482
|||||
Db 483 ATCAGTGGAGGGCAGAAACAGAGGATCGCAATTTGGCTCGTCCCTTAGTTTCGAAACCCCAA 542
|||||
Qy 482 sIleLeuLeuLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 502
|||||
Db 543 GATTTCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTGAGAAGCAAGTACAGCTGTC 602
|||||
Qy 502 InAlaAlaLeuGluLys 507
|||||
Db 603 AAGCTGCACCTGGAGAG 619
|||||

RESULT 5
BM082401
LOCUS BM082401
DEFINITION AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
5', mRNA sequence.
ACCESSION BM082401
VERSION BM082401.1 GI:22274409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2492 row: c column: 15
High quality sequence stop: 677.
FEATURES
source
1..929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/clone_lib="NIH MGC 43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',

```

adaptor: GGCACAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH-MGC Library. "

BASE COUNT 281 a 176 c 237 g 234 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.8e-90 Length: 929  
Score: 926.50 Matches: 188  
Percent Similarity: 79.30% Conservative: 61  
Best Local Similarity: 59.87% Mismatches: 58  
Query Match: 14.41% Indels: 7  
DB: 14 Gaps: 2

US-09-873-409-6 (1-1195) x BQ882401 (1-929)

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThr-Valll 195  
Db 1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGGTGTAT 60  
QY 195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLyl 215  
Db 61 TGCATTGGAGGACAAAGAAAGAACTTGAAGGTACAAACAAAATTTAGAAAGACTAA 120  
QY 215 sAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePh 235  
Db 121 RAGNATTGGATAGAAAGCTATTACAGCAATATTTCTATAGGTGCTTCTCTGCT 180  
QY 235 eMetAsnGlyThrTyrGlyLeuAlaPheTyrThrGlyThrSerLeuIleLeuAsnGlyGl 255  
Db 181 GATCATGTCATCTTATGCTCTGGCCCTCTGGTATGGAGCACCCCTGGTCTCTCAGGGGA 240  
QY 255 uProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSertY 275  
Db 241 A-----TATCTTATGGACAAAGTACTCACTGATTTCTTTCTGTATTAATTTGGGGCTTT 294  
QY 275 rCySileGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPh 295  
Db 295 TAGTGTGGACAGCATCTCCAGCATTTGAGCATTTGCAATGCAAGAGGAGCAGCTTA 354  
QY 295 eHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTy 315  
Db 355 TGAATACTTCAAGATAATTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCA 414  
QY 315 rLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSe 335  
Db 415 CAACACAGATAATATTAAAGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCATC 474  
QY 335 rArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVa 355  
Db 475 TGAAGAAGAGTTAAGATCTTGAAGGCTTGAACCTTGAAGGTGCAGAGTGGGAGAGGTT 534  
QY 355 lAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLe 375  
Db 535 GGCCCTGGTTGGAACAGTGGCTGTGGGAAGACACACAGTCCAGTCCATGATGAGAGGCT 594  
QY 375 uTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVa 395  
Db 595 CTATGACCCACAGAGGGATGCTAGTGTGTGATGACAGGATATTAGGACCAATAATGT 654  
QY 395 lArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrTh 415  
Db 655 AGGTTTCTACGGGAATACTATTGGTGTGGTGTGAGTACGAACTGTATGTTTCCACCAC 714  
QY 415 rIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAl 435  
Db 715 GATAGCTGAAAACATTCCTTATGCGCGTGAATAATGTCACCATGATGATGAGATTGAGAAAGC 774  
QY 435 aAlaArgGluAlaAsnAla-TyrAspPheIleMetGluPheProAsnLysPheAsnThrIl 455  
Db 775 TGTCCAGGAAGCCAAATGCCCTATGACTTTTATCATGAAACTGCCCTCATAAATTTGACACCC 834

QY 455 euValGlyGluLysGlyAlaGlnMetSerGlyGly---GlnLysGlnArgIleAlaIleA 474  
Db 835 TCGTTGGANAAGAAGAGGGCCCAATTGAGTGGTGGGCGCAACAAACCAAGGATCTCCCTTGG 894  
QY 474 lAArg-AlaLeuVal-ArgAsnProLysIleLeu 484  
Db 895 CACGTGGCCCTGGTTCCCAACACCCCAAGATCTC 928  
RESULT 6  
LOCUS AL040762 545 bp mRNA linear EST 29-FEB-2000  
DEFINITION DKFZp343C1815\_r1 434 (synonym: htes3) Homo sapiens cDNA clone  
ACCESSION DKFZp343C1815 5', mRNA sequence.  
VERSION AL040762  
KEYWORDS AL040762.1 GI:5409708  
SOURCE EST.  
ORGANISM human.  
REFERENCE 1 (bases 1 to 545)  
AUTHORS Bloecker H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
Wiemann, S.  
TITLE EST (Bloecker, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Bloecker H  
MIPS  
Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;  
Sponsored by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
s1 sequence also available.  
This clone (DKFZp343C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source 1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZp343C1815"  
/clone\_lib="434 (synonym: htes3)"  
/tissue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

BASE COUNT 171 a 89 c 138 g 147 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.84e-90 Length: 545  
Score: 918.00 Matches: 181  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 15.27% Indels: 0  
DB: 9 Gaps: 0

US-09-873-409-6 (1-1195) x AL040762 (1-545)

QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325  
Db 3 AGTATAGATTAACCTTTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTGGAA 62  
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345  
Db 63 TTTTAAATGTTTCTTCAATTATCCATCAAGACCATCTATCAAGATTCGAAAGGCTCTG 122  
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365  
Db 123 AATCTCAGAAATTAAGTCTGGAGAGACAGTCGCGCTTGGTGGTCTCAATGGCAGTGGGAAG 182

```

QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 183 AGTACGGTAGTCAGCTCTGCGAGAGTTATATGATCCGATGATGGCTTTATCATGGTG 242
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 243 GATCAGATGACATCAGAGCTTTAAATGTGGGCATTATCGAGCCATATTCGAGTGGTT 302
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 303 AGTCAAGAGCGCTGTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGACGAGAT 362
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 363 GATGTGACTGATGAAGAGATGGAGAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 422
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 465
Db 423 ATGGAGTTTCTTAATAATTTAATACATTGGTAGGGGAGGAGGAGGAGGAGGAGGAGGAG 482
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 483 GGGCAGAAACAGAGATCGCAATTCGTGCTGCTAGTTCGAAACCCCAAGATTCGATT 542
QY 486 Leu 486
Db 543 TTA 545

RESULT 7
LOCUS BM471690 948 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
5', mRNA sequence.
ACCESSION BM471690
VERSION BM471690.1 GI:18520732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/PTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12233 row: b column: 14
High quality sequence stop: 569.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5539117"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab host="DHI0B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 285 a 187 c 194 g 282 t
ORIGIN
Alignment Scores: 9.03e-83 Length: 948
Pred. No.:

```

```

Score: 856.50 Matches: 186
Percent Similarity: 89.67% Conservative: 5
Best Local Similarity: 87.32% Mismatches: 14
Query Match: 14.25% Indels: 8
DB: 13 Gaps: 1

US-09-873-409-6 (1-1195) x BM471690 (1-948)
QY 266 ValPhePheSerValIleHisSerTyrCysIleGlyAlaValProHisPheGlu 285
Db 205 GTTTCCTTTAGTGTATATCATCATAGCAGTTATGATGGAGCAGCCTCCTCACTTGA 264
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db 265 ACCTTCGCATAGCCGAGGAGCTGCTTTTCATATTTCCAGGTATTGATAAGAACCC 324
QY 306 SerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGAATCCATAGAGGAACCTGTGGA 384
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db 385 TTTAAATAATGTTTCITTCATATTCATCAAGACCATCTATCAAGATTCTGAAGGCTCG 444
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 445 AATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTGGTCTCAATGGCAGTGGGAAG 504
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 505 AGTACGGTAGTCTGAGCTTCTGAGAGGTTATATGATCCGAGTATGCGGATGCTTATCATGGTGTG 564
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVal 405
Db 565 GATGAGAATGACATCAGAGCTTTAAATGTGGGCATATTCAGACCATATTTGGAGTGGT 624
QY 405 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 424
Db 625 TAGTCAAGAGCCCTGTTTTCGGGACCAACCATCAGTAACAATATCAAGTATGGGAGCG 684
QY 424 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla--AsnAlaTyrAsp 443
Db 685 AGATGATGGGACTGATGAAACAAATGGAACAAACAAAGGGAACCAAGGCGGAATGA 744
QY 444 PheIleMetGln-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 462
Db 745 TTTATCATGGAATTTCCCAAAAAAATTTATCATTTGGAAGGGGAAAAAAGACCTTCA 804
QY 462 nMetSerGlyGlyGlnLysGlnArg 470
Db 805 AGGAATGGAAGGCGCAAAACCCCAAG 829

RESULT 8
LOCUS AK014319 2676 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
enriched library, clone:322401P09;ATP-binding cassette, sub-family
B (MDR/TAP), member 8, full insert sequence.
ACCESSION AK014319
VERSION AK014319.1 GI:12852089
KEYWORDS HTC; CAP trap.
SOURCE Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
Clone:322401P09.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci P. and Hayashizaki Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

```





```
QY 35 ValGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54
|||
|||
|||
Db 579 GTGCTCAGTCCCGTAAGCTCAGCGTCCAGCTGCTCTACTGTAC-----GGTGT 629

QY 55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTyrPheIleThrAlaAlaArg 73
|||
|||
|||
Db 630 CAGGAGCTGCTGACCTTGGATACCTAGTGTCTGCTGCCACATT-----GGTGAGCGC 683

QY 74 GlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnApeIleGlyTyr 93
|||
|||
|||
Db 684 ATGGCCATGACATCGGAAAGCCCTTTTTCAGCTCCCTGCTCGGCAAGACATGCTTTC 743

QY 94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111
|||
|||
|||
Db 744 TTTGATGCCAAAAGACAGAGGCGAGTAGTGTGCTGTGACTACTGTATGTGCAAGATTC 803

QY 112 -----IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124
|||
|||
|||
Db 804 AAGTCATCCTTCAAGCTTGTTCATCTCCAGGAGCTG----- 839

QY 125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValLysGlyTyrLys 143
|||
|||
|||
Db 840 ---CGCAGCTGCACCCAGGATGTGTAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGCG 896

QY 144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSer 163
|||
|||
|||
Db 897 CTACCTGATGTGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 956

QY 164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
|||
|||
|||
Db 957 TCAGGCTCCGAAAGCTGTCTCGCCAGTGTCCAGGAGCATTTGCCAGGCAACAGGTGTA 1016

QY 184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
|||
|||
|||
Db 1017 GCAGATGAGGCCCTTGGCAATGTTCGAGCTGTGCGGCCCTTCCGATGGAGAAGAGGAG 1076

QY 204 LeuGlnArgTyrThrGlnAsnLeuLysAsp-----AlaLysAspPheGlyIle 219
|||
|||
|||
Db 1077 GAGGAACGCTATCAAGCAGAACTGGAGTCATGCTGTGTAAGCAGAACTGGGCG--- 1133

QY 220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe-----PheMetAsn 237
|||
|||
|||
Db 1134 ---AGGGGATCGGCTGTGTTCCAAAGGCTCTCCAAATCGCTTTCAACTGTATGTGCTTG 1190

QY 238 GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu---Pro 256
|||
|||
|||
Db 1191 GGCACC-----CTGTTTCATTTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAA 1238

QY 257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCys 276
|||
|||
|||
Db 1239 GGGGAGACCTCATGCTCTCTGCTGGTCCAGACAGTACAGAGCTTATGGCCAGC 1298

QY 277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly-----AlaAla 294
|||
|||
|||
Db 1299 CTCCTGCTCTGTTTGTGTCAG-----GTGGTACGTGGCTGAGTGCGGGA 1343

QY 295 PheHisIlePheGlnValIleAspLysIlePheProSerIleAspAsnPheSerThrAlaGly 314
|||
|||
|||
Db 1344 GCCCGAGCTTTCGAATACATGCGCCCTGAGCCCTGTCATCCCA-----TTGACCGGGGGC 1397

QY 315 Tyr-----LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 332
|||
|||
|||
Db 1398 TACTGCAATCCCAACAGAGCATTTGGTGTTCATCATCTTCCAAATGTCCACCTTCAGC 1457

QY 333 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 352
|||
|||
|||
Db 1458 TACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGACTTCACCCCTGAAGCTGCCCTCTGCG 1517

QY 353 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 372
|||
|||
|||
Db 1518 AAGATTGTGGCTCTTGTGGGCGAGTCTGGGGAGGAAAGACACAGTGTGCTCTCTGCTG 1577
```

```
QY 373 GlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAla 392
|||
|||
|||
Db 1578 GAACGGCTTCTATGACCTTGAGCTCGGTGAGCTTGGATGGGCGATGACCTCGGAAC 1637

QY 393 LeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGluProValLeu 411
|||
|||
|||
Db 1638 CTCACCCCTCTGCTCGGCGGCGAGGTATAGTGTTCATCAGCCAGGAGCGAGTCTCTG 1697

QY 412 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 431
|||
|||
|||
Db 1698 TTTGCAACACCATCATCGAGATATCCGATTGGGAAGCTGGATGCTTCCGATGAAGAG 1757

QY 432 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 451
|||
|||
|||
Db 1758 GTGTACACAGCTGCACAGAGCAATGCCAGAGTTTCATCAGCAGCTTCCCGATGCG 1817

QY 452 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 471
|||
|||
|||
Db 1818 TACAGCAGCTGTGGTGGTGAGCGGGGCAACCTTGTCTGTGGTGGCCAGAGCGCCCTA 1877

QY 472 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 491
|||
|||
|||
Db 1878 GCCATCGACAGTCCCTCATCAGCAGCCACAGTGTCTGATCTGGACGAGCCACCAT 1937

QY 492 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 511
|||
|||
|||
Db 1938 GCGCTAGATCGAGATCCGAGAGGGTGTACAGAGGCGCTGAGAGCGCCCTGAGCCGCG 1997

QY 512 ArgThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 531
|||
|||
|||
Db 1998 CGCACCGTGTGGTATTGCCACCGGCTTAGTACTGTCCGTGCGTGCAGCCCATCATCAT 2057

QY 532 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 551
|||
|||
|||
Db 2058 GTCATGCCCAATGCCCAAGTCTGTAGGCTGGGAGCCACAGAGAACTCTCTTAAAAAGGC 2117

QY 552 GlyLeuTyrTyrSerLeuValMetSerGlnAspIle 563
|||
|||
|||
Db 2118 GGGCTTATTACAGAGCTTATCCGAGAGCAAAACCTG 2153

RESULT 9
BM904842 998 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
DEFINITION 5', mRNA sequence.
ACCESSION BM904842
VERSION BM904842.1 GI:19355221
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12279 row: f column: 24
High quality sequence stop: 738.
Location/Qualifiers
1..998
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557655"
/clone_lib="NIH_MGC_72"
FEATURES
source
```

/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: Not I;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

BASE COUNT 271 a 223 c 248 g 256 t

ORIGIN

Alignment Scores:

Pred. No.: 2,93e-80 Length: 998  
Score: 834.00 Matches: 164  
Percent Similarity: 74.12% Conservative: 68  
Best Local Similarity: 52.40% Mismatches: 77  
Query Match: 13.87% Indels: 4  
DB: 14 Gaps: 0

US-09-873-409-6 (1-1195) x BM904842 (1-998)

Qy 718 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 737  
Db 61 GATGTGAGTGGTTTGGTATGACCCCTAAACACCACTGGAGCATTGACTACCGCTCGCC 120  
Qy 738 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 757  
Db 121 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTCTGTAATTACCCAGAAT 180  
Qy 758 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleThrGlyTrpGluMetThrPhe 777  
Db 181 ATAGCAAACTTGGGACAGGAATAATATATCTTCACTATGTTGGCACTAACCACTG 240  
Qy 778 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaLaMet 797  
Db 241 TTACTCTTAGCAATTTGACCCATCATTTGCAATAGCAGGAGTTGTTGAAATGTTG 300  
Qy 798 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 817  
Db 301 TCTGGCAAGCACTGAAGAGATAAGAAAGAACTAGAAAGGTGCTGGGAAGATCGCTACTGAA 360  
Qy 818 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 837  
Db 361 GCATAGAAACATTCGAAACCGTGTCTTTGACTCAGGAGCAGCAAGTTGAAACATATG 420  
Qy 838 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 857  
Db 421 TATGCTCAGAGTTTGAGGTACCATACAGAACTCTTTGAGGAAGACACACATCTTGA 480  
Qy 858 SerCysTyAlaPheSerHisAlaPheIleTyAlaThrAlaAlaGlyPheArgPhe 877  
Db 481 ATTACATTTTCTTCCACCCAGGCAATGATGATATTTTCTTATGCTGATGTTTCCGGTT 540  
Qy 878 GlyAlaTyLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 897  
Db 541 GGAGCTTACTTGTGGGCACATAAATCATGAGCTTTGAGGATGTTCTGTAGTATTTCA 600  
Qy 898 AlaIleAlaTyGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTySer 917  
Db 601 GCTGTTGCTTTGGTGCCATGGCGGTGGGCAAGTCAGTTCACTTCTCTGACTATGCC 660  
Qy 918 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 937  
Db 661 AAAGCCAAAATATCAGCAGCCACATCATGATCATTTGAAAAAACCCTTTGATTGAC 720  
Qy 938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957  
Db 721 AGTACAGCAGCAGGAGCCCTAATGCCCAACACATTTGGAAGAAATGTCACTTTGGGTGA 780  
Qy 958 ValSerPhePheTyProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe 977  
Db 781 GTTGATTCNACTATCCACCCAGCCGACATCCAGTGTCTCAGGACTGAGCCCTGGA 840  
Qy 977 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrs 997  
.....

Db 841 GGTGAAGAGGGGCGAGACCCCTGGCTCTGTGGGGCAGCAGTGGCTGTGGGAAGACACAG 900  
Qy 997 ex-ValGlnLeuLeuGlnArgLeuTyArgProValGlnGlnGlyValLeuPheAspGly 1016  
Db 901 TGGGTCCAGCTCTCTGGAGCGGTCTACAAACCCCTTGGCAAGGAATAATGCTGCTTGATGCC 960  
Qy 1017 ValAspAlaLys-GluLeuAsnValGlnTrp 1026  
Db 961 AAAAGAATAAAACCGACTGGATGTTCCATGG 991

RESULT 10

BF584668 BF584668 894 bp mRNA linear EST 12-DEC-2000  
LOCUS 602098406F1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4218385 5',  
DEFINITION mRNA sequence.

ACCESSION BF584668

VERSION BF584668.1 GI:11658386

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 894)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM9798 row: 1 column: 02

High quality sequence stop: 651.

FEATURES

source

1..894

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4218385"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: Not I;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.6 Kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 263 a 203 c 242 g 185 t

ORIGIN

Alignment Scores:

Pred. No.: 1.98e-73 Length: 894  
Score: 771.00 Matches: 171  
Percent Similarity: 79.72% Conservative: 57  
Best Local Similarity: 59.79% Mismatches: 48  
Query Match: 12.82% Indels: 15  
DB: 12 Gaps: 3

US-09-873-409-6 (1-1195) x BF584668 (1-894)

Qy 266 ValPhePheSerValIleHisSerSerTyrcysIleGlyAlaAlaValProHisPheGlu 285

Db 5 GTCCTCTTTTCGTGTATAT-CGAGCATTCAGTTGGACAGGCATCTCCAAATATTGAA 63

Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305

Db 64 GCCTTCGCCAATGCAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCC 123

Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrlsProGluSerIleGluGlyThrValGlu 325

Db 124 AGTATAGACAGTCTTCTCAAGAGTGGGCAACCAACACAGACACATACAGGAATCTGGAA 183

QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysValGlyLeu 345  
 Db 184 TTTAAGATATTACCTTCAGTACCTCCCATCTCGAAGAAGCTTCAGATCTTGAAGGCCCTC 243

QY 346 AsnLeuArgIleLysSerGlyThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365  
 Db 244 AATCTGAAGGTGAAGACGACAGCGTGGCTGGTGGCAACAGTGGCTGTGGAAGA 303

QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385  
 Db 304 AGCACAACTGTCAGCTGTGTCGAAGGCTCTACGACCCCTAGATGCAATGTCAGTATC 363

QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405  
 Db 364 GACGGACAGACATCAGAACATCAATGATGAGGTATCTGAGGAGATCATTTGTGTGGTG 423

QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425  
 Db 424 AGTCAGAACTGTGTGCTGTTTGGCCACCATGTCGCGAGAACATTCGTATGCGCGAGAA 483

QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445  
 Db 484 GATGTCCATTCAGTATGATGATGAGTGTGCAAGGAAGCCATGCTATGACTTTCATC 543

QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 465  
 Db 544 ATGAACCTGCCCCCACTATTTGACACCTGTTGTTGAGAGAGGGGCGAGTGGTGG 603

QY 466 GlyGlnLysGlnArgIleAlaAlaAlaArgAlaLeuValArgAsnProLysIleLeuIle 485  
 Db 604 GGACAGAAACAGAGAAATCGC-ATTGCCCGGGCCCTGTCGCGCAATCCCAAGTGCCTTTC- 661

QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaIle 505  
 Db 662 GTGACGAAGCCACCTCAGC-CTGTGATACAGAAAGTGAAGCTGGGGTTCAGGCGGCACCT 720

QY 505 uGluLysAlaSerLysGlyArgThrThrIle-----ValValAlaHisArgLe 521  
 Db 721 AGATAGGTACAAAGCGGAGGACCACTATGGGAGCTCATCGCTGTATCGCTCGAAGCT 780

QY 521 uSerThrIleArgSerAlaAspLeuIleVal-----ThrLeuLysAspGlyMetLeuAl 539  
 Db 781 GACGCCATCGCGCGGAAGGGGG-CTTGTGTGGAAGCACTCTGAAAGC-----TC 830

QY 539 aGluLysGlyAlaHis 544  
 Db 831 TGAAGGAGGTACTCAT 846

RESULT 11  
 BG248052 1019 bp mRNA linear EST 13-FEB-2001  
 LOCUS 602359987F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG248052  
 VERSION BG248052.1 GI:12757867  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1019)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-t@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10334 row: n column: 21

High quality sequence stop: 650.

## FEATURES

source

Location/Qualifiers  
 1..1019  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_lib="IMAGE:4488404"  
 /clone\_lib="NCI\_CGAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT 276 a 262 c 283 g 196 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,85e-73 Length: 1019  
 Score: 770.50 Matches: 154  
 Percent Similarity: 79.17% Conservative: 36  
 Best Local Similarity: 64.17% Mismatches: 47  
 Query Match: 12.82% Indels: 3  
 DB: 12 Gaps: 1

US-09-873-409-6 (1-1195) x BG248052 (1-1019)

QY 950 GluGlyAsnLeuGluPheArgGluValSerPheTyrProCysArgProAspValPhe 969

Db 4 GAAGGAAATGTGCATATTAGTGGAGTCGTGTTCACATATCCACCCGACCCAGCATCCCA 63

QY 970 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 989

Db 64 GTGCTTCAGGGGTGAGCCTTCAGGTGAAGAGGGCCACGCTGGCCCTGTGGGCGAGC 123

QY 990 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 1009

Db 124 AGTGGCTGGGGAGAGCAGCTGTCAGCTGCTCGAGGGCTTCTACGACCCCATGGCT 183

QY 1010 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1029

Db 184 GGATCAGTGTCTAGATGCGCAAGAAATAAAGCAACTGAATGTCAGTGGCTCCGAGCA 243

QY 1030 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1049

Db 244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTTCTTTGACTGCGAGCATCGCAGAAACATT 303

QY 1050 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1069

Db 304 GCCTACGGAGACAAACGCGGGTGTGTTCTTATGAGAGATTGTGAGGGCAGCCCAAGGAG 363

QY 1070 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1089

Db 364 GCCAACATCCACCATTCATCGACTCGCTACCTGATAAATACAAACACAGTAGGAGAC 423

QY 1090 LysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 1109

Db 424 AAAGGCACCTCAGCTGTGGGTGGGAGAGCAGCGCATGCCATCGCAGCCCTCGTC 483

QY 1110 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnSerGlu 1129

Db 484 AGACAGCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGA 543

QY 1130 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValThr 1149

Db 544 AAGGTTGTCCAGGAAGCGCTGGCAAAAGCAGGAAGGCGCACCTGCATTTGTATCGCT 603

QY 1150 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1169

Db 604 CACCGCTGTCCACCATCCCAAGACCGGGACTTGTATCGGTGGTGAATTCAGAGAACCGGAGG 663

Qy	1169	leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe	1186
Db	664	TCAAGGGAGCAGCGGCACCCCAACAGCAAGTGTTGGTCGAGAAGGGCATCTACTTC	719
RESULT	12		
BG293345			
LOCUS			
DEFINITION	602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5', mRNA sequence.	726 bp mRNA linear EST 21-FEB-2001	
ACCESSION	BG293345		
VERSION	BG293345.1	GI:13052943	
KEYWORDS	EST,		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: The Cepko Laboratory CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LHAM10371 Row: 1 column: 09 High quality sequence stop: 658.		
FEATURES	source location/Qualifiers		
	1..726		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:4502552"		
	/clone_lib="NIH_MGC_94"		
	/tissue_type="retina"		
	/lab_host="DHI0B (phage-resistant)"		
	/note="Organ: eye; Vector: pCMV-SPORT6; Site:1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
BASE COUNT	189 a 220 c 190 g 127 t		
ORIGIN			
Alignment Scores:			
Pred No.:	1.28e-71	Length:	726
Score:	753.00	Matches:	152
Percent Similarity:	84.29%	Conservative:	25
Best Local Similarity:	72.38%	Mismatches:	30
Query Match:	12.52%	Indels:	3
DB:	12	Gaps:	0
US-09-873-409-6 (1-1195) x BG293345 (1-726)			
Qy	984	ValAlaPheValGlySerSetGlyCysGlyLyserThrservalGlnLeuLeuGlnArg	1003
Db	2	CTNGGCCCTGTGGGCAGCAGTGGCTGCGGAAGAGCACAGTGGTCCAGTGTCTGAGCGC	61
Qy	1004	LeutyAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaIysGluLeuAsn	1023
Db	62	TTCATATGACCCCATGGCTGGATCATGTGCTTTAGATGGTCAAGAAGCAAAGAACTCAAT	121
Qy	1024	ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys	1043
Db	122	GTCCAGTGGCTCCGAGCTCACTGGGCATTGTGTCCCAGGAACCATCTCTTTGACTGC	181
Qy	1044	SerIleAlaGluAsnIleAlaTyrglyAspAsnSerArqValValProLeuAspGluIle	1063
Db	182	AGCATCGCAGAGAAACATGCCCTATGGACAACAGCCGGTCTGTGCCTCATGATGAGATT	241



[illegible]



```
/strain="C57BL/6J"
/db_xref="FANTOM DB:9230106F14"
/db_xref="MGD:MGI:1911565"
/db_xref="taxon:10090"
/clone="9230106F14"
/sex="male"
/tissue_type="epididymis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
86..-872
/notes="data source:SPRR, source key:P21447, evidence:ISS
putative
similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
3) (MDR1A)"
/codon_start=1
/protein_id="BAB32070.1"
/db_xref="GI:12860873"
/db_xref="MGD:MGI:1924956"
/translation="MANSEKNGLEQNTQRYGFLQEQVPGVGNQAVGPIEIRFADNL
DIIVLTGLIASINGATVPLMSLVGLSEIDSLINGCLVQTNRTKYQNCSTQEKLINE
DIIVLTLYIGIGAAALIFGYQISPMVITAAROTTRIRKOFHSLAODISWFDGSD
ICELNTRMTGDKLDCGIGDKIPLMFQNISGFSIGLIVSLIKSWKLSLVLSSTPLI
MASSALCSRMISLISKELDVSKAGAVAEALSSIQVTAFGA"
BASE COUNT      262 a 178 c 187 g 245 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:      1.53e-70      Length:      872
Score:          744.50      Matches:      151
Percent Similarity: 80.37%      Conservative: 21
Best Local Similarity: 70.56%      Mismatches: 27
Query Match:     12.38%      Indels:      15
DB:              11          Gaps:        2

US-09-873-409-6 (1-1195) x AK020318 (1-872)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 230 ATGACTCTGGGAATATTAGCATCCATGATAAATGAGCCCGCTCCTTTAAATGTCCTG 289
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGATGCGCTAGTACAAACTAACAGAACT 349
Qy 41 SerPhe-----PheArgLeuThr 46
Db 350 AAATATCAGAACTGTTCTCAGACTCAGAAAAGCTGAATGAAGATATCATTTGTTGACT 409
Qy 47 LeuTyrTrpValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 410 CTATATTATATTGGATAGGAGCAGCTGCCCTCATTTTGGCTATGTACAGATTTCTCTC 469
Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 470 TGGGTCACTAGTCAGCCGCGCAACACACAGAATCCGAAAACAGTTTTTTTCATTCAATT 529
Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 530 TTGGCAAGACATCAGCTGGTTTGTATGGCAGTGCATCTGTGAACCTTAACACCCGCATG 589
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 590 ACTGGTGACATCAACAACTCTGTGTGATTTGGAGATAAGATCCCTCTGATGTTTCAG 649
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db 650 AACATATCTGGGTTTCTATTGGCCTGGTGATAAGTTTGTATAAAAAGCTGGAAACTCTCC 709
Qy 146 LeuValThrLeuSerThrProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
Db 710 CTGGTGGTTCTGTCTACATCTCCTCATATGGCTTCATCGGCACGTGCTCTAGGATG 769
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
```

```
Db 770 ATTATCTCATTACCAGCAAGGAGCTGGATGCTATTCCAAAGCTGGGCTGGCTGAA 829
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
Db 830 GAAGCCTTGTCTCATCATCAATCCAAACAGTCACAGCCTTTGGAGCC 871
```

Search completed: March 31, 2003, 13:51:24  
Job time : 8245.55 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 175.002 Seconds  
(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-6  
Perfect score: 6012  
Sequence: 1 MILGILASLVNGACPLMPL.....QELLNRDIYFKLVNNAQSVQ 1195

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09873409/runat\_27032003.115422.19259/app\_query.fasta\_1.7544  
-DB=Issued Patents NA -QPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1.1.578 @runat\_27032003.115422.19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTCUS COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3545	59.0	4646	1 US-08-181-471-2	Sequence 2, Appli
2	3545	59.0	6505	2 US-08-793-610-5	Sequence 5, Appli
3	3545	59.0	9318	2 US-08-793-610-6	Sequence 6, Appli
4	3543.5	58.9	4264	2 US-08-784-649A-1	Sequence 1, Appli
5	3543.5	58.9	4264	2 US-08-784-649A-5	Sequence 5, Appli
6	3543	58.9	4669	6 US06352-3	Patent No. 5206352
7	3527	58.7	4669	2 US-08-752-447-1	Sequence 1, Appli
8	3527	58.7	4669	4 US-09-316-167-1	Sequence 1, Appli
9	3517.5	58.5	4233	3 US-09-120-513-1	Sequence 1, Appli
10	3517.5	58.5	4233	4 US-09-450-105-1	Sequence 1, Appli
11	3497	58.2	4669	2 US-08-583-276-18	Sequence 18, Appli
12	2282.5	38.0	4002	2 US-08-996-545-1	Sequence 1, Appli

13	2282.5	38.0	4002	2 US-08-996-545-3	Sequence 3, Appli
14	2282.5	38.0	4002	4 US-09-328-320-1	Sequence 1, Appli
15	2282.5	38.0	4002	4 US-09-328-320-3	Sequence 3, Appli
16	2280	37.9	2726	1 US-08-461-823-1	Sequence 1, Appli
17	2273.5	37.8	4047	2 US-08-612-734B-1	Sequence 1, Appli
18	2211	36.8	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	2166	36.0	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	2116	35.2	4224	1 US-08-612-521-1	Sequence 1, Appli
21	1731.5	28.8	6143	1 US-08-612-521-3	Sequence 3, Appli
22	1703	28.3	3924	2 US-08-996-644-3	Sequence 3, Appli
23	1703	28.3	3924	3 US-09-352-552-3	Sequence 3, Appli
24	1703	28.3	3927	2 US-08-996-644-1	Sequence 1, Appli
25	1703	28.3	3927	3 US-09-352-552-1	Sequence 1, Appli
26	1664	27.7	3909	1 US-08-232-537-1	Sequence 1, Appli
27	1659.5	27.6	3792	4 US-09-351-224E-10	Sequence 10, Appli
28	1525	25.4	3999	4 US-09-351-224E-9	Sequence 9, Appli
29	1109	18.4	13188	4 US-08-961-527-70	Sequence 70, Appli
30	1057	17.6	7760	4 US-08-961-527-63	Sequence 63, Appli
31	996.5	16.6	441529	4 US-09-103-840A-1	Sequence 1, Appli
32	994.5	16.5	4403765	4 US-09-103-840A-2	Sequence 2, Appli
33	945.5	15.7	7186	4 US-08-961-527-39	Sequence 39, Appli
34	938	15.6	6492	4 US-08-961-527-188	Sequence 188, Appli
35	761.5	12.7	2376	1 US-08-394-880B-1	Sequence 1, Appli
36	718	11.9	1749	4 US-09-134-001C-1893	Sequence 1893, Appli
37	698	11.6	2061	4 US-09-061-764A-17	Sequence 17, Appli
38	687.5	11.4	5889	1 US-08-463-092B-5	Sequence 5, Appli
39	687.5	11.4	5889	2 US-08-462-109A-5	Sequence 5, Appli
40	687.5	11.4	5889	2 US-08-460-907B-5	Sequence 5, Appli
41	687.5	11.4	5889	3 US-08-463-179A-5	Sequence 5, Appli
42	687.5	11.4	5889	3 US-08-461-384B-5	Sequence 5, Appli
43	674.5	11.2	5120	3 US-08-772-270A-6	Sequence 6, Appli
44	674.5	11.2	8370	2 US-08-488-706-1	Sequence 1, Appli
45	671.5	11.2	5011	1 US-08-463-092B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-08-181-471-2  
; Sequence 2, Application US/08181471  
; Patent No. 5641508  
; GENERAL INFORMATION:  
; APPLICANT: Li, Lingna  
; APPLICANT: Lishko, Valeryi K.  
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL COMPOSITIONS TO HAIR FOLLICLES  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Thomas Fitting  
; STREET: 12526 High Bluff Drive, Suite 300  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92130  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/181.471  
; FILING DATE: 13-JAN-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/041,553  
; FILING DATE: 02-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: ANT0029P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-792-3680  
; TELEFAX: 619-792-8477

## ; INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 4646 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHEetical: NO
;   ANTI-SENSE: NO
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 425..4267
; US-08-181-471-2
```

## Alignment Scores:

```
Pred. No.: 0 Length: 4646
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 1 Gaps: 9
```

US-09-873-409-6 (1-1195) x US-08-181-471-2 (1-4646)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGAAATTTAGAAATGATCGATGTC 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 695 AACATCACTAATAGAGTATCAATGATACAGGGTCTTCATGAACTGGAGGAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 755 ATGACCAAGGTATGCCTATTATTACAGTGGAAATGGTCTGGGGTGGTGGTCTGCTTAC 814
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGGAAGACAAATACACAAAATTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATAATGCGACAGGATAGGCTGGTTGATGTCACGATGTTGGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGACAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAC 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGGAAGCTTAACCTTGATTTTGGCCATCAGTCTGCTGTTCTGGACATGTCAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1115 GTCTGGGCAAGATACCTATCTTCAATTTACTGATAAAGAACTCTTAGGCTATGCCAAGCT 1174
QY 181 GlyAlaValAlaGluLeuValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTGCTATTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 1235 AAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGCTAAAGAAATTTGGGATTAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
```

```
DB 1295 AAAGCTATTACAGCCAAATATTTCTATAGCTGCTGCTTCTCTGCTATGCTATGCTATT 1354
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 1355 GCTCTGGCCCTTCTGGTATGGGACCACCTTGGTCTCTCAGGGGAA-----TATTCTATT 1408
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 1409 GGACAGTACTCACTGTATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 1468
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 1469 TCTCCAAGCATTGAAGCATTTCGAATTCAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 1529 ATTGATAATAAGCAAGTATTGACAGCTATTGAGAGGTGGGACAAACAGATAATATT 1588
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1648
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1649 ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGACAGCGTGGCCCTGGTGGAAAC 1708
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 1709 AGTGGCTGTGGGAAGACACACAGCTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1768
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1769 GGGATGGTCAGTGTGATGGACAGGATATTAGGACCATAAATGTAAAGTTCTACGGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1829 ATCATTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1888
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1889 CGCTATGGCCGTAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAGGCAAT 1948
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
DB 1949 GCCTATGACTTTATCATGAACTGCCTCATAAATTTGACACCTGCTGGTGGAGAGAGGG 2008
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 2009 GCCCAGTTGAGTGGTGGGAGAGCAGAGGATCGCCATTTGCAGCTGCCCTGGTTCGCAAC 2068
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 2069 CCCAAGATCTCTGCTGGTGGGACAGCTGAGCTGAGCTTGGACACAGAAAGCGAGAGTG 2128
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyValArgThrThrIleValValAlaHisArg 520
DB 2129 GTTCAGTGGCTCTGGATAAGGCCAGAAAAGGTGGGACACCATTTGTGATGCTCATCGT 2188
QY 521 LeuSerThrIleArgSerAlaPheLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
DB 2189 TTGCTACAGTTCGTAAATGCTGACGTCATCGCTGGTTCATGATGAGTGCATGTTGGAG 2248
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
DB 2249 AAGGAAATCATGATGAATCATGAAAGAGAAAGGAGGATTTTACTTCAAACTTGTCACAATG 2308
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
DB 2309 CAGACAGCAGGAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAT 2368
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589
```

```
Db 2369 GATGCTTGGAAATGCTTCAAAATGATTCAGATCCAGTCTAATAAGAAAAAGATCAACT 2428
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluIleSerThrGlnSerLys 607
Db 2429 CGTAGAGTGTCCGTGGATCACAAGCCCAAGACAAAGCTTAGTACCAAGAGGCTCTG 2488
Qy 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTTP 627
Db 2489 GATGAAGTATATACCTCCAGTTCCTTTGGAGGATATGAAGCTAAATTTAACTGAATGG 2548
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db 2549 CCTATTGTTGTTGGTGTATTTGTCGCAATATATAATGGAGGCTTGCACACAGCATTT 2608
Qy 648 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db 2609 GCAATAATATTTCAAGATATAGGGGTTTTTACAAGAAATTCATGATCCTGAAACAAA 2668
Qy 667 LysHisAspAlaGluIleTySerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db 2669 CGACAGATAGTAACCTGTTTTCACATATTGTTCTAGCCCTTGGAAATTTATTTCTTTATT 2728
Qy 687 SerTyPheMetGlnGlyLeuPheTyGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db 2729 ACATTTTCTTCCAGGGTTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGGCTC 2788
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyGlnAspIleAlaIlePheAspGluLysGlu 726
Db 2789 CGATACATGGTTTTCGATCCATGCTCAGACAGAGTGTGATGGTTGATGACCCCTAAA 2848
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db 2849 AACACCCTGGAGCATTTGACTACAGGCTCGCCCAATGATGCTCTCAAGTTAAAGGGCT 2908
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db 2909 ATAGGTTCCAGGCTTCTGTGTAATTTACCCGAATATAGCAAAATCTTGGGACAGGAAT 2968
Qy 767 IleSerPheIleTyGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db 2969 ATATCCTTCACTATGTTGGCAACTAACACTGTTACTCTAGCAATTTGACCCATCATT 3028
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db 3029 GCAATAGCAGGAGTGTGTAATGAAATGTTCTCTGGACAACGACCTGAAAGATAAGAAA 3088
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db 3089 GAACTAGAAGGTCTGGGAAGATCGCTACTGGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyTrpGluMetLeuGlnThrGlnHis 846
Db 3149 TCTTTCACTCAGGACGAGAAGTTTGAACATATATGATGCTCAGAGTTTGCAGGTACCATAC 3208
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyTrpAlaPheSerHisAlaPhe 866
Db 3209 AGAAACTCTTTGAGGAAGACACACATCTTTTGGAAATACATTTTCTTCCACCGGCAATG 3268
Qy 867 IleTyPheAlaTyTrpAlaAlaGlyPheArgPheGlyAlaTyTrpLeuIleGlnAlaGlyArg 886
Db 3269 ATGTATTTTCTTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGTGGGCACATAAATC 3328
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyTrpGlyAlaMetAlaIle 906
Db 3329 ATAGCTTTGAGGATGTTCTGTAGTATTTTCAGCTGTGTCTTGTGGTCCATGGCCGCTG 3388
Qy 907 GlyLysThrLeuValLeuAlaProGluTyTrpSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db 3389 GGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAGCCAAATATACAGAGCCACATC 3448
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysPro 946
Db 3449 ATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTACAGCAGGAGGCTTAATGCGC 3508
```

```
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyTrpProCysArgPro 966
Db 3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCACCCGACCG 3568
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db 3569 GACATCCCACTGCTTTCAGGGACTGAGCTGGAGGTGAAGAGGGCCAGACGCTGGCTCTG 3628
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuArgLeuTyTrpAsp 1006
Db 3629 GTGGCAGCAGGTGGCTGTGGGAAGACACAGTGTCCAGCTCTCTGGAGCGGTTCACGAC 3688
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTTP 1026
Db 3689 CCCTGGCAGGGAAGTGTCTTGTATGGCAAAAGAAATAAAGCAGCTGAATGTTCACTGG 3748
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db 3749 CTCGAGACACCTCGGCATCGTGTCCAGAGGCCATCTCTGTTGACTGCAGCATTTGCT 3808
Qy 1047 GluAsnIleAlaTyTrpGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db 3809 GAGAATTTGCTTATGGAGACAAACAGCCGGTGTGTACAGGAAGAGATCTGTGAGGGA 3868
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyTrpAsnThrGln 1086
Db 3869 GCAAGGAGGCCAACATACATACCTTCATCGAGTCACTGCCTTAATAAATATAGCACTAAA 3928
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db 3929 GTAGAGACAAAGGAACCTCAGCTCTCTGTGGTGGCAAAACAACGCAATGCCATAGCTCGT 3988
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db 3989 GCCCTTGTAGACAGCCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db 4049 GAAAGTGAAGAGTTGTCCCAAGAGCCCTGGACAAAGCCAGAGAGGCCGCACTGCATT 4108
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 4109 GTGATGTCTCAGCGCTGTCCACCATCCAGAATGCCAGATTAAATAGTGTGTTCAGAAT 4168
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyTrpPhe 1186
Db 4169 GGCAGAGTCAAGAGCATGGCAGCATCAGAGCTGCTGGCACAGAAAGGCACTCTATTTT 4228
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 4229 TCAATGGTCAAGTCCAGGCT 4249
```

## RESULT 2

```
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/793,610
/ FILING DATE: 07-MAR-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: DE P 44 31 973.8
/ FILING DATE: 08-SEP-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: DE 195 03 952.1
/ FILING DATE: 07-FEB-1995
/ APPLICATION DATA:
/ APPLICATION NUMBER: PCT/BP95/03175
/ FILING DATE: 10-AUG-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Berman, Richard J.
/ REGISTRATION NUMBER: 39,105
/ REFERENCE/DOCKET NUMBER: P1614-7007
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 638-5000
/ TELEFAX: (202) 638-4810
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 6505 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA
/ US-08-793-610-5

```

## Alignment Scores:

```

Pred. No.: 0 Length: 6505
Score: 3545.00 Matches: 681
Percent Similarity: 75.55% Conservative: 246
Best Local Similarity: 55.50% Mismatches: 264
Query Match: 58.97% Indels: 36
DB: 2 Gaps: 9

```

US-09-873-409-6 (1-1195) x US-08-793-610-5 (1-6505)

```

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 1967 ATGGTGGGAACTTTGGCTGGCATCATCCATGGGCTGGACTTCCTCATCATGCTG 2086
QY 21 ValIleGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 2077 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGAAATTTAGAGATCTGATGTC 2086
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhe----- 43
Db 2087 AACATCACTAATAGAGTGTATCATATCATACAGGGTTCTTCATGAATCTGGAGGAGAC 2146
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 2147 ATGACCAAGGTATGCTATTATTACAGTGAATTTGGTCTGGGCTGGTGTGCTGCTTAC 2206
QY 62 IleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrIlysergileArgIlyGln 81
Db 2207 ATTCAGGTTTCATTTTGGTGGCTGGAGAGTGGAGACAAATACACAAATTTAGAAACAG 2266
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 2267 TTTTTCATGCTATATGACAGAGGATAGGCTGTTGATGTGCACGATGTGGGGAG 2326
QY 102 LeuAsnThrArgMetThr-----AspIleAspIlySerAspGlyIleGlyAspIlyle 120
Db 2327 CTTAAACACCCGACTTACAGATGATGCTCTTAAGATTAATCAAGTTATTGGTGACAAAAT 2386
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 2387 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCATCTGGGTTTATAGTAGGATTTACACGT 2446

```

```

QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 2447 GGTGGAGAGCTAACCTTTGCTGATTTTGGCCATCAGTCTCTGTTCTTGGAGTGTGAGTGT 2506
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 2507 GTCTGGGCAAGATACTATCTTCATTCTGATTAAGAACTCTTAGCGTATCAAAAGCT 2566
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 2567 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGACTGTGATTCATTTGGAGACAA 2626
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 2627 AAGAAAGAACTTGAAGAGGTACAAACAAATTTAGAAGAACTTAAAGAAATTTGGATAAG 2686
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 2687 AAAGCTATTACAGCCAATATTTCTATAGTGTGCTTCTCTGCTGCTGATCTATCGATCTTAT 2746
QY 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 2747 GCTCTGGCCCTTCTGGTATGGGACACCTTGGTCTCTCAGGGGAA-----TATTCATAT 2800
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 2801 GCACAAGTACTCACTGATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 2860
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 2861 TCTCCAAGCATTTGAAGCATTTTGCAATTCAGAGGAGCAGCTTATGAAATCTTCAAGATA 2920
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 2921 ATGTATATTAACCAAGTATTTCAGACTATTTCAGAGGTGGGACACAAACACATATATAT 2980
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 2981 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 3040
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 3041 ATCTTGAAGGGCTGAACCTGAAGGTGACAGTGGGACACGGTGGCCCTGTTGGTGAAC 3100
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 380
Db 3101 AGTGGCTGTGGAGAGGACACACAGTCCAGCTGATGCAGAGCTCTATGACCCCCACAGAG 3160
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 3161 GGGATGCTCAGTGTGATGACAGAGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3220
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 3221 ATCATTTGGTGTGGTGTGAGTGTGATTTGTTTGGCCACCAAGTGTGATGAGTGTGAGTGT 3280
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 3281 CGCTATGGCCGTGAATATGTACCATGATGATGATGATGATGATGATGATGATGATGATGAT 3340
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 3341 GCCTATGACTTTTATCATGAACTGCTCTATAAATTTGACACCTTGTGTTGAGAGAGAGGG 3400
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 3401 GCCCAGTTGAGTGTGGGAGAGGAGGATCGCAATTTGCACTGCGCTGCTGCTGCTGCTGCT 3460
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 3461 CCCAAGATCTCTGCTGATGAGGCCAGCTGACCTTGGACACAGAAAGCCAGAGAGTG 3520
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520

```

Db 3521 GTTCAGGTGGCTCTGGATGAAGCCAGAAAAGGTTCGGACACCATTTGTGATAGCTCATCGT 3580  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 3581 TTGCTACAGTTGCTGAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGGGAG 3640  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrSerLeuValMetSer 560  
Db 3641 AAAGGAATCATGATGAATCATGAAAGAGAAAGGCAATTTACTTCAAACTTTGTCACAATG 3700  
Qy 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
Db 3701 CAGACAGCAGGAATGAAGTTGAATTAGAAAATGCGAGCTGATCAATCCAAAAGTGAATTT 3760  
Qy 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
Db 3761 GATGCTTCGGAATGCTTCAATAGATTCAAGATCCAGTCTTAATAGAAAAGATCAACT 3820  
Qy 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607  
Db 3821 CGTAGGAGTTCGCTGGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3880  
Qy 608 GluIleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGluTyr 627  
Db 3881 GATGAAAGTATACCTCCAGTTCCTTTGCGAGGATTATGAAGCTAAATTTAACTGAATCG 3940  
Qy 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
Db 3941 CCTATTGTTGTTGGTGTATTTGTCGCAATTAATAATGGAGGCTGCAACGAGCATTT 4000  
Qy 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 4001 GCAATAATATTTCAAGATTATAGGGTTTTTACAAGAAATTGATGATCCTGAAACAAA 4060  
Qy 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
Db 4061 CGACAGATAGTAACCTGTTTTCACATATTGTTCTAGCCCTTGGAAATTATTTCTTTATT 4120  
Qy 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706  
Db 4121 ACATTTTCCCTCAGGGTTTCATTTTGGCAAGCTGGAGAGATCTCTACCAAGCGGCTC 4180  
Qy 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleAlaIlePheAspGluLysGlu 726  
Db 4181 CGATACATGGTTTTCATCCATGCTCAGACAGATGTGAGTTGGTTGTATGACCCCTAAA 4240  
Qy 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
Db 4241 AACACCACTGGAGCATGACTACAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGCT 4300  
Qy 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
Db 4301 ATAGGTTCCAGGCTCTGTAATATCCCAAGATATAGCAAAATCTTGGGACAGGAATAAT 4360  
Qy 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786  
Db 4361 ATATCCTTTATCTATGTTGGCAACTAACACTGTTACTCTAGCAATTGTACCCATCAT 4420  
Qy 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
Db 4421 GCAATAGCAGGAGTGTGAAATGAAATGTTGCTGCGACAAGCACTGAAAGATAAGAAA 4480  
Qy 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
Db 4481 GAACTAGAAGGCTCTGGGAAGATCGCTACTGAAAGCAATAGAAAACCTTCCGAACCGTTGT 4540  
Qy 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846  
Db 4541 TCTTTGACTCAGGAGCAGAACTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 4600  
Qy 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866

Db 4601 AGAAACTCTTTGAGGAAAGCACACATCTTTTGAATATACATTTTCTTACCAGGCAATG 4660  
Qy 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886  
Db 4661 ATGTATTTTTCCTATGCTGGATGTTTCCGTTTGGAGCCTACTTGTGTGCACATAAATC 4720  
Qy 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
Db 4721 ATGAGCTTTGAGGATGTTCTGTGTAGTATTTTCACTGTGTCTTGTGGTCCATGGCCGTG 4780  
Qy 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926  
Db 4781 GGGCAAGTCAGTTTCATTTGCTCTGCTGATATGCCAAAGCCAAATAATATCAGACGCCACATC 4840  
Qy 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946  
Db 4841 ATCATGATCATTTGAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCTTAATGCCG 4900  
Qy 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
Db 4901 AACACATTTGGAAGGAATGTACATTTGGTGAAGTTGTATTCAACTATATCCACCCGACCG 4960  
Qy 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986  
Db 4961 GATATCCCGAGTCTTCAGGAGCTGGAGGTGGAAGAGGCGGCGGCTGGCTCTG 5020  
Qy 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006  
Db 5021 GTGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCTCTGGAGCGGTCTACGAC 5080  
Qy 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026  
Db 5081 CCCTTGGCAGGGAAGTGTCTGTGATGTCAAAGAAATAAAGCGACTGTAATGTTCAGTGG 5140  
Qy 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046  
Db 5141 CTCGAGCAGACCTGGGCATCTGTCGCCAGGAGCCCATCTCTGTTCAGCTGCAGCATGTCT 5200  
Qy 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066  
Db 5201 GAGACATTTGCTATGGAGACAAACAGCCGGTGGTGTGCACAGGAAGAGATCGTGAGGGCA 5260  
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086  
Db 5261 GCAAGGAGGCGCAACATACATACATGCTTCATCGAGTCACTGCCTAATAAATATATGACATAA 5320  
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 1106  
Db 5321 GTAGGAGACAAAGGAACTCAGCTCTCTGTGTGCGCAAAACAAACGCAATGCCATAGCTCGT 5380  
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 5381 GCCTTGTAGACGCTCATATTTGCTTTTGGATGAAGCCACGCTCAGCTCTGGATACA 5440  
Qy 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
Db 5441 GAAAGTGAAGAGGTTGTCCAAGAGCCCTGGCAAAAGCCAGAGAGCGCGCACCTGCATT 5500  
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValLeuHisAsn 1166  
Db 5501 GTGATGCTCAGCGGCTGTCCACCATCCAGAAATGCGAGACTTAATAGTGTGTTCAGAAAT 5560  
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnLeuLeuArgAsnArgAspIleTyrPhe 1186  
Db 5561 GGCAGAGTCAAGAGCATGGCAGCATCAGAGCTGCTGGCACAGAAAGGCATCTATTTT 5620  
Qy 1187 LysLeuValAlaGlnSer 1193  
Db 5621 TCAATGCTCAGTGTCCAGCT 5641

RESULT 3

US-08-793-610-6

; Sequence 6, Application US/08793610





QY 421 LysTyrGlyArgAspValThrAspGluMetGluAlaAlaArgGluAlaAsn 440  
Db 3240 CGTATGGCGTGAATGTCCACCATGTGAGATTGAGAAAGCTTCAGAAAGCAAT 3299  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
Db 3300 GCTATGACTTTATCATGAACCTGCTCATTAATTTGACACCCCTGGTTGGAGAGAGGG 3359  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAlaArgAlaLeuValArgAsn 480  
Db 3360 GCCCAGTTGAGTGGCGAGAGAGAGATCGCCATTCGACGTGCGCTTGGTTCGCAAC 3419  
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaIleAspSerGluSerLysSerAla 500  
Db 3420 CCCAAGATCCTCTGCTGGATGAGGCCACGCTGACCTTGGACACAGAAAGCGAAGCATG 3479  
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaAlaHisArg 520  
Db 3480 GTTCAGTGGCTCTGATAGAGCCAGAAAGGTGCGACCACTTGTGATGATGCTCATCGT 3539  
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 3540 TTGCTACAGTTGCTAATGCTGACGTCATCGCTGTTTCGATGATGGAGTCATTGGGAG 3599  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
Db 3600 AAGGAATCATGATGAATCATGAAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 3659  
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
Db 3660 CAGACAGCAGAAATGAAGTTGAATTAGAAATGACGCTGATGAATCCAAAGTAGTAAAT 3719  
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
Db 3720 GATGCTTGGAAATGCTTCAATGATTCAGATCCAGTCTAATAGAAAGAAAGATCAACT 3779  
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607  
Db 3780 CGTAGGAGTCCGCTGGATCACAAGGCCAAGACAGAAAGCTTAGTACCAAGAGGCTCTG 3839  
QY 608 GluIleSerLeuProGluValSerLeuLysIleLysIleLysLeuLysProGluTyr 627  
Db 3840 GATGAAGTATACCTCCAGTTCTCTTTGGAGGATATGAAAGCTAAATTTAACTGAATGG 3899  
QY 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
Db 3900 CCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3959  
QY 648 SerIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
Db 3960 GCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAAATTCATGATCCTGAAACAAA 4019  
QY 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
Db 4020 CGACAGAATAGTAACTGTTTTCACATGTTTCTAGCCCTTGGAAATATTTCTTTTATT 4079  
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLeuIleLeuThrMetArgLeu 706  
Db 4080 ACATTTTCTCTCAGGGTTTCATTTGGCAAGAGCTGGAGAGATCTCCACCAAGCGGCTC 4139  
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIlePheAspGluLysGlu 726  
Db 4140 CGATACATGGTTTCCGATCCATGCTCAGACAGGATGTTGGTTGGTTGATGACCTAAA 4199  
QY 727 AsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
Db 4200 AACACCACTGGACATGACTACGAGCTCGCAATGATGCTCTCAAGTTAAAGGGCT 4259  
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
Db 4260 ATAGTTTCCAGGCTCTGCTGAATTTACCAAGATATAGCAATCTCTGGACAGGAATAT 4319  
QY 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786

Db 4320 ATATCCTTCTATCTATGTTGGCAACTTAACACTGTACTCTTAGCAATTGTACCATCAT 4379  
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAlaAsnLysAspLysGln 806  
Db 4380 GCATAGCAGGAGTGTGTAATGAAATGTTGTCTGGACAGCACTGAAGATAGAAA 4439  
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
Db 4440 GAACTAGAAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCGCAACCGTTGT 4499  
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 846  
Db 4500 TCTTTGACTCAGAGCAGAAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4559  
QY 847 ArgAsnThrSerLysLysAlaGlnIleGlySerCysTyrAlaPheSerHisAlaPhe 866  
Db 4560 AGNAATCTTTGAGGNAAGCACACATCTTTGGNAATTACATTTCTTCCACCCAGGCAATG 4619  
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886  
Db 4620 ATGTATTTTCTATGCTGATGTTTCCGTTTGGAGCTACTTGGTGGCACATAAATC 4679  
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
Db 4680 ATCAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTTCTTGGTCCCATGGCGTG 4739  
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926  
Db 4740 GGGCAGTCACTGTTCTCTGCTATGCCAAGCCAAATATATCAGACGCCCAATC 4799  
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluLysLysPro 946  
Db 4800 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGSCCTAATGCCG 4859  
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
Db 4860 AACACATTCGAGAGAAATGTCATTTGTTGAGTTGTATTCAACTATCCCAACCCGACCG 4919  
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986  
Db 4920 GACATCCAGTCTTCAGGAGCTGAGCTGGAGGTGAAAGGCCAGACGCTGGCTCTG 4979  
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp 1006  
Db 4980 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGTCTCAGCAC 5039  
QY 1007 ProValGlnGlyClnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026  
Db 5040 CCTTTGGCAGGGAAGTGTCTGTGTCGCAAGAAATAAAGCGACTGAATGTTCAAGTGG 5099  
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValIlePheAsnCysSerIleAla 1046  
Db 5100 CTCGAGACACCTGGGCATCGTCCAGGAGCCCATCTCTGTTTACCTGACATGCTGT 5159  
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066  
Db 5160 GAGAACATTCCTATGGACACACAGCCGGTGTGTCTCACAGGAAGAGATCGTGAGGGCA 5219  
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086  
Db 5220 GCAAGAGAGGCCAACATACATCGCTTCATCGAGTCACTGCCTTAATAAATATAGCACTAA 5279  
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlnLysGlnArgLeuAlaIleAlaArg 1106  
Db 5280 GTAGGAGACAAAGGAACCTCAGCTCTCTGTGTGGCCAGAAACAACGCAATTCCTAGTCTGT 5339  
QY 1107 AlaLeuLeuGlnLysProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 5340 GCCCTGTTAGACGCTCATATTTGCTTTTGGATGAAGCCACGCTCAGCTCTGGATACA 5399  
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
:::|||||

```

Db 5400 GAAAGTGAAGAGGTGTCACAGAGCCCTGGACAAAGCCAGAGAGCCGCGACCTGCATT 5455
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db 5460 GTGATTGCTCACCGCTGTCCACCAATCCAGAAATGAGACTTAATAGTGGTGTTCAGAAT 5519
Qy 1167 GlyValIleValGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db 5520 GGCAGAGTCAAGAGAGATCGCACGATCAGCAGCTGCTGGCCACAGAAAGGCATCTATTTT 5579
Qy 1187 LysLeuValAsnAlaGlnSer 1193
Db 5580 TCAATGCTCAGTGTCCAGGCT 5600

RESULT 4
US-08-784-649A-1
; Sequence 1. Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.: 0 Length: 4264
Score: 3543.50 Matches: 682
Percent Similarity: 75.63% Conservative: 246
Best Local Similarity: 55.58% Mismatches: 262
Query Match: 58.94% Indels: 37
DB: 2 Gaps: 10

US-09-873-409-6 (1-1195) x US-08-784-649A-1 (1-4264)
Qy 1 MetLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 292 ATGCTGGTGGAACTTGGCTGCCATCCATGGGCTGGAGCTTCCTCTCATGATGCTG 351
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 352 GTGTTTTCGAGAAATGACAGATATCTTTGCAAAATGAGGAAATTTAGAAGATCTGATGTCA 411

```

QY 381 GlyPheIleMetValAspGluAsnAspIleAraGlaLeuAsnValArgHisTyrArgAsp 400  
DB 1483 GGGATGTCAGTGTGATGGACAGAGATATATAGACCAATAAATGTAAAGGTTTCTACGGGAA 1542  
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnile 420  
DB 1543 ATCAATTGGTGGTGGAGTCAGGAACCTGTAATGTTTGGCCACACGATAGCTGAAACATTT 1602  
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
DB 1603 CGCTATGGCCGTGAAATGTCCACATGGATGAGTTGAGAAAGCTGTCAAGGAAGCCAAT 1662  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460  
DB 1663 GGCTATGACATTTATCATGAACTGCCTCAATAATTTTGACACCCCTGGTTGGAGAGAGGG 1722  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
DB 1723 GCCAGTTGAGTGGTGGGACAGAGAGATCGCCATTTGCACGTCCTGGTTCCGAAC 1782  
QY 481 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
DB 1783 CCCAAGATCCTCTGCTGATGAGGCCACGTGAGCCTTGGACACAGAAAGCGAAGCAGTG 1842  
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
DB 1843 GTTCAGGTGGCTCTGGATAGGCCAGAAAAGGTTCGACCAACCATTTGTGATGATCATCGT 1902  
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
DB 1903 TTGTTCTACAGTTGCTGAATGCTGACGTCATCGCTGGTTTCGATGATGGATCATTTGGGAG 1962  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
DB 1963 AAAGGAATCATGATGAATCATGAAAGAGAAAGGCAATTTACTTCAAACTTGTCACAATG 2022  
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
DB 2023 CACACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2082  
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
DB 2083 GATGCCCTTGGAAATGCTTCAATAGATTCAAGATCCAGTCTAATAAGAAAAGATCAACT 2142  
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607  
DB 2143 CGTAGAGTGTCCGTGGATCAGACAGCCAGACAGAAAGCTTAGTACCAAGAGGCTCTG 2202  
QY 608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 627  
DB 2203 GATGAAAGTATACCTCCAGTTTCCCTTTTGGAGGATTTATGAAGCTAAATTTAACTGAATGG 2262  
QY 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
DB 2263 CCTATTTTGTGGTGTATTTTGTGCATTTAATAATGGAGGCTTCAACCAAGCATTTT 2322  
QY 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
DB 2323 GCATAATATATTTCAAGATTTATAGGGTTTTTACAGAAATGTGATGATCTCGAACAATA 2382  
QY 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
DB 2383 CGACAGAAATAGTAACATTTGTTTTCACATTTGTTTCTAGCCCTTGGAAATTTTCTTTTATT 2442  
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706  
DB 2443 ACATTTTCTTCAGGGTTTCAATTTTGGCAAGCTTGGAGAGATCTCTCAACAGCGGCTC 2502  
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 726  
DB 2503 CGATACATGGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTGATGACCCCTAAA 2562  
QY 727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746

DB 2563 AACACACCTGGAGCATTTGACTACCAAGGCTCGCCAAATGATGCTGCTCAAGTTAAAGGGGT 2622  
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
DB 2623 ATAGGTTCCAGGCTTGTGTAATTTACCCAGAAATATAGCAAAATCTTGGGACAGAAATAT 2682  
QY 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786  
DB 2683 ATATCCTTTCATCTATGTTGGCAACTAAACACTGTACTCTTAGCAATTTGACCCATCAT 2742  
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
DB 2743 GCATAGCAGGAGTTGTTGAAATGAAATGTTGCTCGGAACAAGCACTGAAGATATAAGAA 2802  
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
DB 2803 GAACTAGAAGGTCCTGGGAGATCCGTACTGAAGCAATAGAAAACCTTCGAAACCGTTGT 2862  
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846  
DB 2863 TCTTTGACTCAGGAGCAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 2922  
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866  
DB 2923 AGAAACTCTTTGAGGAAACACACATCTTTGGAATTTACATTTTCTTCCACCGGCAATG 2982  
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886  
DB 2983 ATGATATTTTCTATGCTGGATGTTTCGGTTTGGAGCTACTTGTGGCACATAAATCTC 3042  
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
DB 3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTTGTTGTTGTTGTTGTTG 3102  
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926  
DB 3103 GGGCAAGTCAGTTTCAATTTGCTCTGCTATGCAAGCCAAATATATCAGCAGCCCAATC 3162  
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysPro 946  
DB 3163 ATCATGATCATTTGAAATAAACCCCTTTGATTGACAGTACAGCAGGAGGCTTAATGGCG 3222  
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
DB 3223 AACACATTTGGAAGAAATGTCATTTTGTGTAAGTTGTTATTCAACTATCCACCCGACCG 3282  
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986  
DB 3283 GACATCCCACTGCTTTCAGGGACTGAGCCTGGAGGTGAAGAGGGCCAGCGCTGGCTCTG 3342  
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006  
DB 3343 GTGGGACAGATGGCTGTGGGAAGACACATGGTGGTCCAGCTCTGAGCGGTTCTACGAC 3402  
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026  
DB 3403 CCCTTTGGCAGGAAAGTGTCTGCTGATGSCAAAGAAATAAAGCGACTGAATGTTCACTG 3462  
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046  
DB 3463 CTCCGAGCACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGACGATGCT 3522  
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 1066  
DB 3523 GAGAACATTTGCTATGGAGACAAACAGCGGGGTGTCTCAGGAAGAGATTTGTGAGGGA 3582  
QY 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086  
DB 3583 GCAAGAGGAGCCAAACATACATGCTTTCATCGAGTCACTGCTTAATAATATAGACTAAA 3642  
QY 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106

Db 3643 GTAGGACAAAGAACTCAGCTCTCTGGTGGCCAGAAAACAGCATTGCCATAGCTCGT 3702  
QY 1107 AlaleuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 3703 GCCTTGTTAGACAGCTCATATTTTGGTTTGGATGAAGCCAGCTCAGCTCTGGATACA 3762  
QY 1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
Db 3763 GAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGAGCCGACCTGCATT 3822  
QY 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166  
Db 3823 GTGATGCTCACGCCCTGTCCACATCCAGATGCGACTTAATAGTGGTGTTCAGAAAT 3882  
QY 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186  
Db 3883 GGCAGAGTCAAGAGAGCATGCGACGATCAGCAGCTGCTGCGACAGAAAGCACTATTTT 3942  
QY 1187 LysLeuValAsnAlaGlnSer 1193  
Db 3943 TCAATGCTCAGTCCAGGCT 3963

## RESULT 5

US-08-784-649A-5  
; Sequence 5, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; APPLICANT: Chen, Gang  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,649A  
; FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: Reg.No. 5830697 36,677  
REFERENCE/DOCKET NUMBER: 06037/007001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4264 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-784-649A-5  
Alignment Scores:  
Pred. No.: 0 Length: 4264  
Score: 3543.50 Matches: 682  
Percent Similarity: 75.63% Conservative: 246  
Best Local Similarity: 55.58% Mismatches: 262  
Query Match: 37 Indels: 37  
DB: 2 Gaps: 10

US-09-873-409-6 (1-1195) x US-08-784-649A-5 (1-4264)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20  
Db 292 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGCTGGAGCTTCCTCTCATGATGCTG 351  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 352 GTGTTTGGGAAATGACAGATATCTTGGCAATGTCAGGAAATTTAGAAATCTGATGTCA 411  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 412 ACATCATCTAATAGAGTGCATATCAATGATACAGGTTCTTTCATGAATCTGGAGGAAGAC 471  
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 472 ATGACCAGGTATGCCTTATTATACAGTGAATTTGGTGGGCTGCTGTTGCTGCTTAC 531  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 532 ATTCAAGTTTCAATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 591  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 592 TTTTTCATGCTATATATGACAGGAGATAGGCTGTTTGTGTCACCATGTTGGGAG 651  
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspDlyIleGlyAspLysIle 120  
Db 652 CITTAACACCCGACTTACAGATGATGTCCTCAAGATTAATGAAGGAATTTGGTGACAAAT 711  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 712 GGA-----ATGTTCCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTACAC 768  
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 769 GGTGGAAGCTTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTCAGCTGCT 828  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 829 GTCTGGCAAGACATATCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 889 GGAGCAGTAGCTCAGAGAGCTTTGGCAGCAATTAGAACCTGTGATTCATTTGGAGACAA 948  
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 949 AGAAAGAACTTGAAAGGTACACAAATAATTTAGAAAGAGCTAAAGAAATTTGGATTAAG 1008  
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 1009 AAAGCTATTACAGCCAATATTCTATAGTGTGCTTCTTCTGCTGATCTATGATCTTAT 1068  
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 1069 GCTCTGGCCTTCTGTTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1122  
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 1123 GGACAAGTACTCACTGTATTCTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 1182  
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValaAlaPheHisIlePheGlnVal 300  
Db 1183 TCTCCAAGCATTGAAGCATTTCGAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1242  
QY 301 IleAspLysGlyProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1243 ATTGATAATAAGCAAGTATTGACAGCTATTGACAGTATTCAGAGTGGGCAACCAAGATAAT 1302  
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1303 AAGGGAAATTTGGAAATTCAGAAATGTTTCACTTCACTTCCATCTCGAAAGAGTTAAG 1362

QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
DB 1363 ATCTTTGAAGGGCTGAACCTGAAGCTGCAGATGGGCACAGCGTGGCCCTGGTTGGAAAC 1422  
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
DB 1423 AGTGGCTGGGAAGAGCAACAGCTGCAGCTGATGACAGAGGCTATATGACCCACAGAG 1482  
QY 381 GlyPheIleMetValAspGluAsnAspIleAraGlaLeuAsnValArgHisTyrArgAsp 400  
DB 1483 GGGATGGTCAGTGGTGGACAGAGATATAGGACCAATAATGTAAGGTTTCTACGGGAA 1542  
QY 401 HisIleGlyValIleSerGlnProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
DB 1543 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1602  
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
DB 1603 CGCTATGGCCGGTGAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
DB 1663 GCCTATGACTTTATCATGAACCTGCCTCATAAATTTGCACCCCTGGTTGGAGAGAGGG 1722  
QY 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
DB 1723 GCCCAGTTGAGTGGTGGGAGAGCAGAGATCGCCATTGTCACGTCCTCGTTGTCGAAC 1782  
QY 481 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
DB 1783 CCCAAGATCCTCTGCTGATGAGCCAGCTCAGCCTTGACACAGAAAGCAAGCAGTG 1842  
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
DB 1843 GTTCAGTGGCTCTGTAAGGCCAGAAAGGTCGGACCCACTTGTGATGCTCATCGT 1902  
QY 521 LeuSerThrIleArgSerAlaLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
DB 1903 TTGCTACAGTTTCGTAAATGTCAGCTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1962  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
DB 1963 AAGGAAATCATGATGAACCTCATGAAGAGAAAGGCAATTTACTTCAAACTTGTCACAA 2022  
QY 561 Gln-----AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574  
DB 2023 CACACAGCAGGAATGAAGTTGAATAGAAAATGCAGCTGATGAATCCAAAAGTGAAAT 2082  
QY 575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal----- 589  
DB 2083 GATGCTTGGAAATGCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2142  
QY 590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluSerThrGlnSerLys 607  
DB 2143 CGTAGGAGTGTCCGTGGATCACAAAGCCCAAGACAGCAAGAGCTTAGTACCAAGAGGCTGT 2202  
QY 608 GlnLeuSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr 627  
DB 2203 GATGAAGATATACCTCCAGTTTCTTTTGGAGGATATGAGCTAAATTTAACTGATGG 2262  
QY 628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647  
DB 2263 CCTATTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2322  
QY 648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666  
DB 2323 GCAATAATATTTTCAAGAGATATAGGGGTTTTCACAGAAATTCATGATCCTGGAACAAA 2382  
QY 667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686  
DB 2383 CGACAGATAGTAACTGTGTTTCTACTATTGTTTCTAGCCCTTGGAAATATTTCTTTATT 2442  
QY 687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyLeuIleLeuThrMetArgLeu 706

DB 2443 ACATTTTCTTCAGGGTTTCACATTTGGCAAAGCTGGAGATCTCTACCAAGGGGCTC 2502  
QY 707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu 726  
DB 2503 CGATACATGGTTTCCGATCCATCTCAGACAGGATGTGAGTTGGTTGATGACCCATAA 2562  
QY 727 AsnSerThrGlyGlyLeuThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746  
DB 2563 AACACCATGGAGCAATGACTACACAGGCTGCCAATGATGCTGCTCAAGTTAAAGGGGT 2622  
QY 747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766  
DB 2623 ATAGTTTCCAGGCTTGTGTAATTAACCCAGATATAGCAATCTTTGGGACAGAAATAT 2682  
QY 767 IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786  
DB 2683 ATATCCTTCATCTATGGTTGGCAACTAACACCTGTTACTCTTAGCAATGTTACCCATC 2742  
QY 787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806  
DB 2743 GCATAGCAGGAGTGTGTTGAAATGAAATGTTGCTCTGGCAACAGCACTGAAGATAAGA 2802  
QY 807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826  
DB 2803 GAACTAGAAGGTGCTGGGAAGATCGCTCTGAAGCAATAGAAAACCTTCCGACCGTTGT 2862  
QY 827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis 846  
DB 2863 TCTTTGACTCAGGACAGAAAGTTTGAACATATGATGCTCAGAGTTTGCAGGTACCATAC 2922  
QY 847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866  
DB 2923 AGAAATCTCTTTCAGGAAAGCACACATCTTTGGAATTTACATTTTCTCCACCCAGGCA 2982  
QY 867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886  
DB 2983 ATGTATTTTCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGTGGCACATAAATCTC 3042  
QY 887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906  
DB 3043 ATGAGCTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTTGTTGTTGTTGTTGTTGTTG 3102  
QY 907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926  
DB 3103 GGCAAGTCAGTTCATTTGCTCTGCTACTATGCCAAAGCCAAATATATCAGACCCACATC 3162  
QY 927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946  
DB 3163 ATCATGATCATTTGAAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTTAATG 3222  
QY 947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966  
DB 3223 AACACATGGAAGAAATGTCATTTGTTGTAAGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3282  
QY 967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986  
DB 3283 GACATCCAGTGTCTCAGGACTGAGCCTGGAGGTGAAGAGGGCCAGACGCTGGCTCTG 3342  
QY 987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnIleLeuGlnArgLeuTyrAsp 1006  
DB 3343 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGGTTCTACGAC 3402  
QY 1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026  
DB 3403 CCTTGGCAGGGAAGTGTGTTGATGTCGAAAGAAATAAGCGCACTGAATGTTAGTGG 3462  
QY 1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046  
DB 3463 CTCGAGACACCTGGGCATCGTGTCCAGGAGCCCATCTCTGTTGACTGCAGCATGCT 3522  
QY 1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066

Db 3523 GAGAACATTGCCTATGTGAGACAAACAGCCGGTGGTGTGTACAGGAAGAGATTGTGAGGCA 3582  
Qy 1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGluLeuProGluLysTyrAsnThrGln 1086  
Db 3583 GCAAGAGAGCCCAACATACATGCTTTCATCGAGTCACTGCCTAATAAATATAGCACTAA 3642  
Qy 1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaAlaArg 1106  
Db 3643 GTAGGACAAAGAACTCAGCTCTCTGTGGCCAGAAACACGCAATGCCATGCTCGT 3702  
Qy 1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126  
Db 3703 GCCCTTGTAGACGCCCTCATATTTTCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762  
Qy 1127 AspSerGlyValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146  
Db 3763 GAAAGTGAAGAGGTGTCCAAAGAACCCCTGGACAAAGCCAGAGAGCCGCACTGCAT 3822  
Qy 1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166  
Db 3823 GTGATTGCTCACCGCTGTCCACCATCCAGATCCAGACTTAATAGTGGTGTTCAGAA 3882  
Qy 1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186  
Db 3883 GGCAGAGTCAAGAGAGTGGCAGCATCAGCAGCTGTGGCAGAGAAAGGCATCTATTT 3942  
Qy 1187 LysLeuValAsnAlaGlnSer 1193  
Db 3943 TCAATGTCAGTGTCCAGGCT 3963

RESULT 6  
5206352-3  
Patent No. 5206352  
Applicant: Roninson, Igor B.; Pastan Ira H.; Gottesman, Michael M.  
Title of Invention: COMPOSITIONS FOR CLONES CONTAINING DNA  
SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
Number of Sequences: 4  
Current Application Data:  
Application Number: US/07/622,836  
Filing Date: 24-SEP-1990  
Prior Application Data:  
Application Number: 892,575  
Filing Date: 01-AUG-1986  
Application Number: 845,610  
Filing Date: 28-MAR-1986  
Seq ID No: 3  
Length: 4669  
5206352-3

Alignment Scores:  
Pred. No.: 0 Length: 4669  
Score: 3543.00 Matches: 681  
Percent Similarity: 75.47% Conservative: 245  
Best Local Similarity: 55.50% Mismatches: 265  
Query Match: 58.93% Indels: 36  
Gaps: 6  
DB:

US-09-873-409-6 (1-1195) x 5206352-3 (1-4669)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 575 ATGGTGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGCA 694  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 695 AACATCTAATAGAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754  
Qy 44 -----ArgLeuThrLeuTyrTyraValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 755 ATGACCAGGTATGCCTATTATTATACAGTGAATTTGGTGTGGGTGCTGGTGTGCTTAC 814  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 815 ATTCAGTTTCTATTTTGGTGCCTGGCAGCTGGAAGCAAAATACACAAATTTAGAAACAG 874  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 875 TTTTTCATGCTATATATGCGACAGAGATAGGCTGGTTGATGTGCACGATGTGGGAG 934  
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATCAAGTTATTGGTGACAAAAT 994  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 995 GGAATGTTCTTTCAGTCAATGGCACTTTTTCACGTGGGTATTATAGTAGGATTTACAC 1054  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 1055 GGTGGAGCTAACCTTGTGATTGTGGCCATCAGTCTCTGTCTTGGACTGTTCAGTGTCT 1114  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 1115 GTCTGGGCAAGATACATACTATCTTCAATTTACTGATAAAGAACTCTTAGCTGATGCA 1174  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 1175 GGAGCAGTAGCTGAAGAGGTGTGGCAGCAATTAGAACCTGTGATTGCATTGTGAGACAA 1234  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 1235 AGAAGAACTTGAAGAGTACACAAATAATTTAGAAAGCTTAAAGAAATTTGGATAG 1294  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 1295 AAAGCTATTACAGCCCAATTTCTATAGTGTCTCTTCTCTGCTGATCTATGATCTTAT 1354  
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 1355 GCTCTGGCTTCTGGTATGGACCACTTGGTCTCTCTCAGGGGAA-----TATTCTATT 1408  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 1409 GGACAGTAGTACTACTGTATCTTTCTGTATTAATTTGGGGCTTTTAGTTGACAGGCA 1468  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1469 TCTCCAAGCATTTGAAGCAATTTGCAAAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 1528  
Qy 301 IleAspLysIleProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1529 ATTGATAAATGAAGCAAGTATTGACAGCTATTCGAAGAGTGGGCAACACAGATAATATT 1588  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1589 AAGGGAATTTGGAATTCAGAAATGTTCACTTACCTTCCATCTCGAAAAGAGTTAAG 1648  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCGACAGCGTGGCCCTGGTTGGAAC 1708  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1709 AGTGGCTGTGGGAAGAGCACACAGCTCAGCTGATCAGAGGCTCTATGACCCACAGAG 1768  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1769 GGAATGTCAGTGTTCATGGACAGGATATTAGACCAATAATGTAAGGTTTCTACGGGAA 1828  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420









Db	2669	CGACAGATAGTAACTGTTTTCACATATGTTTCTAGCCCTTGGAAATTAATTTCTTTTATT	2712
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2729	ACATTTTTCCTTCAGGGTTTCACATTTTGGCAAAAGCTGGAGAGATCCTCACCAAGCGGCTC	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu	726
Db	2789	CGATACATGTTTTCGATCCCATGCTCAGACAGSAGTGTGAGTTGGTTTCATGACCCCTAAA	2848
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2849	AACACCACTGGACATTGACTACAGGCTCGCCAAATGATGTCTCAAGTTTAAAGGGCT	2908
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	2909	ATAGTTTCCAGGCTGCTGTAATTATCCCAAGATATAGCAAAATCTTGGCACAGGAATAAT	2968
Qy	767	IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATAATCCTTCATCATGGTTGGCAACTTAACACTGTTTACTCTTTAGCAATTTGTACCCATCAT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3029	GCAATACAGAGTGTGTGAATGAAATGTTTGTGGACAGCACTGAAAGATAGAAA	3088
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3089	GAACCTAGAAGGCTCTGGGAAGATCGCTACTGAAGCAATAGAAAACCTCCGAACCGTTGTT	3148
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	3149	TCCTTTGACTCAGGACGAGAAGTTTGAACATATGTATGCTCAGAGTTTCAGCGTACCATAC	3208
Qy	847	ArgAsnThrSerLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	3209	AGAAACTCTTTGGAGGAAGCACACATCTTTGGAAATACATTTTCTTCCACCGAGCAATG	3268
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	3269	ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCTACTTGTGGGCACATAAACTC	3328
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	3329	ATGAGCTTTCAGGATGTTCTGTAGTATTTTCAGCTGTTGTCTTTTGGTGCATGCGCGTG	3388
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
Db	3389	GGGCAAGTCAGTTCATTTGCTCTCTGACTATGCCAAGCCAAAATATCAGCAGCCACATC	3448
Qy	927	PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	946
Db	3449	ATCATGATCATTTGAANAACCCCTTTGATTTGACAGCTACAGCAGCGAGGCTTAATCGC	3508
Qy	947	AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro	966
Db	3509	AACACATTTGAAGGAAATGTGCATTTGTTGAAGTTGTATTCAACTATCCACCCGACCG	3568
Qy	967	AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe	986
Db	3569	GACATCCAGTGCTTTCAGGACCTGACCCCTGGAGGTGAAGAGGGCCACAGCTGCTCTG	3628
Qy	987	ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuGlnArgLeuTyrAsp	1006
Db	3629	GTGGGACAGTAGTGCTGTGGAGAGACACAGTGTCTCAGCTCCTGGAGGGTTCTACGAC	3688
Qy	1007	ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp	1026
Db	3689	CCCTTGGCAGGAAAGTGCTGCTTGATGGCAAGAAATAAAGCCGACTGAAATGTTTCAGTGG	3748
Qy	1027	LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla	1046
Db	3749	CTCCGACACCACTGGGCATCGTGTCCGAGAGCCCATCTCTGTTGACTGAGCATGCT	3808



Db	2369	GATGCTCTGGAAATGCTCTTCAAAATGATTCAGAGTCAGATCCAGTCTTAATAGAAAAAGATCAACT	2422
Qy	590	---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys	607
Db	2429	CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAGAGGCGCTCG	2488
Qy	608	GluLeSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTirp	627
Db	2489	GATGAAATAGTACTCCAGTTCCTTTTGGAGGATATGAAGCTAAATTTAACTGAAATGG	2548
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2549	CCTTATTTTGTGTGTGTATTTTGTGCATTAATAATGGAGGCTCGAACACGACATTT	2608
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu	666
Db	2609	GCAATAATATTTCAAAGATTATAGGGTTTTTCAAGAAATTGATGATCCTGAAACAAA	2668
Qy	667	LysHisAspAlaGluLeuTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2669	CGACAGAAATAGTAACCTGTTTTCACATATTGTTTCTAGGCCCTTGGAAATATTCTTTTATT	2728
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2729	ACATTTTTCTCTCAGGGTTCACATTTGGCAAGCTGGAGAGATCCTCACCAGCGGCTC	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTirpPheAspGluLysGlu	726
Db	2789	CGATACATGGTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCATAA	2848
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2849	AACACCACCTGGAGCAATGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTTAAAGGGCT	2908
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	765
Db	2909	ATAGGTTCCAGGCTGCTGTAAATATCCCAAGATATAGCAAAATCTTGGGACAGGAATAAT	2968
Qy	767	IleSerPheIleTyrGlyTirpGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATATCCTTCATCTATGGTTGGCACTAAACACTGTACTCTTAGCAATTTGTACCCATCAT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3029	GCAATAGCAGGAGTTGTGAAATGAAATGTTGTGTGGCAAGCACTGAAAGATAAGAAA	3088
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3089	GAACTAGAAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTCCGAAACCGTGTGT	3148
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluMetLeuGlnThrGlnHis	846
Db	3149	TCCTTTGACTCAGGACGAGAAGTTTGAACATATGATGTCTCAGAGTTTCGAGGTACCATAC	3208
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	3209	AGAAACTCTTTGAGGAAGACACACATCTTTTGAATTTACATTTTCTCACCACCGCAATG	3268
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886
Db	3269	ATGTATTTTTTCTATGCTGATGTTTCCGGTTTGGAGCCTACTTGTGGGCACATAAACCTC	3328
Qy	887	MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle	906
Db	3329	ATGAGCTTTGAGGATGTCTGTTAGTATTTTCAGCTGTGTCTTTTGGTGCCATGCCGTG	3388
Qy	907	GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu	926
Db	3389	GGGCAAGTCAGTTTCATTGCTCTCTGACTATGCCAAAGCCAAAATATCAGCAGCCACATC	3448
Qy	927	PheAlaLeuLeuLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro	946
Db	3449	ATCATGATCATTTGAAAAAACCCCTTTGATTGATGACGTACAGCAGCGAAGGCGCTAATGCG	3508



Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 710 ATGTGGGCAAGTACTGACTTCTATTAATAAGGAACCTCCAGGCTTTATCGGAAAGCT 769  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 770 GGAGCAGTTGCCGAAGAAGTCTTACAGCCATCAGAACTGTGATTCGGTTTGGAGACAA 829  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 830 AAGAAGGAACCTGAAGGTACAAATAAATAATTAGAAGAGCTAAAGAGTTGGCATAAG 889  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 890 AAGCCATCAGCGCCCAACATCTCCATAGTATTGGCTACCTGTTGGTCTATCGCTTAT 949  
Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyThrIle 260  
Db 950 GCACCTGGCAATCTGGTATGGGACCTCCTTGGTCCCTCAATGAA-----TATTCATAT 1003  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 1004 GGACAAGTGTCTACCGTCTTCTCTATTTATTGGGACCTTTCAGTATTGCACATTTA 1063  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1064 GCCCAACATAGAAGCCTTTCGAATGCAAGAGGGCAGCCTATGAATCTTCAAGATA 1123  
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1124 ATTTGATAATGACCAAGCAATTCACAGCTTCTCAACCAAGGGACCAACACAGACATATA 1183  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1184 ATGGGAATTTGGAAATTTAAATGTTTACTTCACTACCCATCAGAACTGAAGTTAAG 1243  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1244 ATCTTGAAGGGCTCAACTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1304 AGTGGCTGTGGGAAAGCAACTGTCTCAGCTGTCTCAGAGGCTCTACACCCCATAGAG 1363  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1364 GCGCAGGTCTGATCAGCGACAGGACATCAGACCATCAATGTGAGGTATCTCGGGAA 1423  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420  
Db 1424 ATCAITGGGGTGTGAGTCAGAAACCGTGTGTTTGGCCACACAGGATTCGCCGAAACATT 1483  
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1484 CGCTATGGCCGAAACCTCACCATGGATGATAGATAAGAAAGCTGTCAAGGAAGCCAT 1543  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
Db 1544 GCGTATGACTTATCATGATGAATCGCCCAAAATTTAAACCCCTGTTGGTGGAGAGGG 1603  
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGATCGCCATTCGCCGGCCCTGGTCCGCAAC 1663  
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1664 CCCAAGATCCTTTGTTGGATGAGGCCAGCTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723  
Qy 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
Db 1724 GTTCAGGCCGCTCTGATAGGCTAGAGAGGGCCGACCAACATTTGATAGCTACCCGC 1783  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540

Db 1784 TTGCTCAGTGGCGAATGCTGACGTCAATGCTGTTTGTGTTGGTGGTCAATGTGGAG 1843  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
Db 1844 CAAGGAATCATGAAGAGCTCATGAAGAGAAAGGCACTTACTTCAAACTTGTCAATGACA 1903  
Qy 561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580  
Db 1904 CAG---ACTAGGAAATGAAATTCAGACAGAAATAATGCTTATGAATCCCAAGTGCAC 1960  
Qy 581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600  
Db 1961 ACTGTGCTCTGAGTTGACTTTCAGAAATCA---AAATCTCTTTTAAATAGGAGATCA 2017  
Qy 601 -----GluGluSerThrGlnSerLysGlu--- 608  
Db 2018 ATTCGACAGAGTATCCACAGAAGCAACAGCAGGAGAGAGACTTAGTTCGAAAGAGGAT 2077  
Qy 609 -----IleSerLeuProGluValSerLeuLysIleLeuLysLeuAsnLysProGlu 626  
Db 2078 GTGGATGAAGATGTGCTTATGCTTTCCTTTTGGCAGATCCTAAAGCTAAATATTAGTGAA 2137  
Qy 627 TrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal 646  
Db 2138 TGCCCTATTAGTTGTGGGTGACTTGTGCTGTTATAAATGGGTGCATACAACCCAGTG 2197  
Qy 647 PheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAspLysThrThr 665  
Db 2198 TTTGCCATAGTGTTCGAAGATTGTAGGGTTTTTCAAGAGACGACGACCATGAAC 2257  
Qy 666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685  
Db 2258 AAACAACGGAATGTAACTTGTGTTTCCCTCTCTCTTCTGTCATGGAATGATTTCTTT 2317  
Qy 686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705  
Db 2318 GTTACGTACTTCTTTCAGAGCTTCACATTTGGCAAGCTGGAGAGATCTCCACCAAGCGA 2377  
Qy 706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLys 725  
Db 2378 CTCCCATACATGCTCTTCAATCCATGCTCGCAGAGATATAAGCTGGTTGATGACCAT 2437  
Qy 726 GluAsnSerThrGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745  
Db 2438 AAAAACACCACTGGCTCGCTGACTACCAAGGCTCGTAGTGACGCTCTAATGTTAAAGGG 2497  
Qy 746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765  
Db 2498 GCTATGGGCTCCAGGCTTGTGATTTACCCAGAAATGTAGCAAAACCTTGGCACAGAAAT 2557  
Qy 766 IleIleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProVal 785  
Db 2558 ATCTATCTTCTAGTCTATGGCTGGCAGCTTACCTTTACTTGTAGTAAATATACCATCTC 2617  
Qy 786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805  
Db 2618 ATTGCTTGGGTGGAATTTATTGAAATGAAACTGTTGCTGCTGCTCAAGCCTTGAAGGACAAG 2677  
Qy 806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 825  
Db 2678 AAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTTGAAAACTTCCGCACCTGT 2737  
Qy 826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845  
Db 2738 GTCTCTTTCATAGGAGCAGAAAGTTTGAACATGATGATGCCAGAGCTTGCAGATACCA 2797  
Qy 846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865  
Db 2798 TACAGAAATGCTTTGAGAAAGACACACGCTTCTTGGGATCACCTTCGCTTCCCCAGGCC 2857  
Qy 866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrIleLeuIleGlnAlaGly 885







Db 830 AAGAGGAACTTGAAGGTTACAAATAAAATTTAGAAAGAGCTTAAAGAGAGTTGGCATAAAG 889  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAenGlyThrTyr 240  
Db 890 AAGCCATCAGCGCCAAACATTTCCATAGTATGCTACTGTTGGTCTATGCGTCTAT 949  
Qy 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAenGlyGluProGlyThrIle 260  
Db 950 GCACGTGCATTTCTGGTATGGGACCTCTGGTCTCTCAATGAA-----TATTCTATT 1003  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaIle 280  
Db 1004 GGACAGTCTTACCCTCTCTCTATTTATTGGGACCTTTTCAGTATGGCAATTTA 1063  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1064 GCCCAACATAGACCTTTGCAATGCAAGGCGGCGAGCTTATGAATCTTCAAGATA 1123  
Qy 301 IleAspLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1124 ATTGATAATGAGCAAGCATTTGACAGCTTCTCAACCAAGGACACAAACAGACAGTATA 1183  
Qy 321 GluGlyThrValGluPheLysAenValSerPheAenTyrProSerArgProSerIleLys 340  
Db 1184 ATGGGAAATTTGGAATTTAAATGTTTACTTCAACTACCCCATCACGAAAGTGAAGTTAAG 1243  
Qy 341 IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1244 ATCTTGAAGGGCTCAACCTGAAGGTGAAGCGGCGACAGCTAGCCCTGGTTGGCAAC 1303  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1304 AGTGGCTGTGGGAAAGACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363  
Qy 381 GlyPheIleMetValAspGluAenAspIleArgAlaLeuAenValArgHisTyrArgAsp 400  
Db 1364 GCGCAGGTGAGTATCGACGACAGCAGCATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAenAenIle 420  
Db 1424 ATCATTTGGGGTGTGAGTCAGGAACCGTGTCTGTTGCCACACAGATTGCCGAAACATT 1483  
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1484 CGCTATGGCCGAGAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543  
Qy 441 AlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyGluLysGly 460  
Db 1544 GCCTATGACTTTCATCATGAACTGCCCCCAAAATTTAAACACCCCTGGTTGGTGAAGAGGG 1603  
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1604 GCGCAGCTGAGTGGGGACAGAAACAGAGATCGCCATTTGCCCGGGCCCTGGTCCCAAC 1663  
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1664 CCCAAGATCTTTGTTGGATGAGCCACGTCAGCCTTGGACACAGAAAGCGGCGGTG 1723  
Qy 501 ValGlnAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
Db 1724 GTTCAGCGCGCTCTGGATAGGCTAGAGAGGCGGACCACTTGTGATGATGCTACCGC 1783  
Qy 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
Db 1784 TTGCTACAGTCGCAATGTGACGTCATTTGCTGGTTTGGATGGTGTGTCATTGTGGAG 1843  
Qy 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 1560  
Db 1844 CAAGGAATCATGAAGGCTCATGAAGAGAGAGGCACTTACTTCAAACTTGTTCATGACA 1903  
Qy 561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580

Db 1904 CAG---ACTAGAGGAATGAAATTTGAACCAAGGAATAATGCTTATGAATCCCAAGTGAC 1960  
Qy 581 ThrAenSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600  
Db 1961 ACTGGTGCCTCTGAGTTGACTTTCAGAAAAATCA---AAATCTCCTTTAATAAGGAGATCA 2017  
Qy 601 -----GluGluSerThrGlnSerLysGlu--- 608  
Db 2018 ATTCGACAGAGTATCCACAGAAAGCAAGACCAAGAGAGAGACTTAGTTTCGAAAGAGGAT 2077  
Qy 609 -----IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAenLysProGlu 626  
Db 2078 GTGGATGAAGATGTCCTATGTTTCTTTTGGCAGATCTTAAGCTAAATATTAGTGAA 2137  
Qy 627 TrpProPheValValLeuGlyThrLeuAlaSerValLeuAenGlyThrValHisProVal 646  
Db 2138 TGGCCCTATTAGTTTGGGTGCTACTTGTGCTGTATATAATGGGTGCATACAAACAGTG 2197  
Qy 647 PheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAspLysThrThr 665  
Db 2198 TTTGCCATAGTCTTTTCAAGATTTGAGGGGTTTTTTCAGAGACGACGACCATGAAACC 2257  
Qy 666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685  
Db 2258 AAACAACGGAATTTGAATCTTTCCTTCTTCTGTCATGGGAATTTCTTTT 2317  
Qy 686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705  
Db 2318 GTTACGTACTTCTTCAAGGCTTCATTTTGGCAAGCTGGAGAGATCTCCACCAAGCGA 2377  
Qy 706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaIleAlaIlePheAspGluLys 725  
Db 2378 CTCCGATACATGCTTCAATCCATGCTGCAGCAGATATAAGCTGGTTGTATGACCAT 2437  
Qy 726 GluAenSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745  
Db 2438 AAAAACCCACCTGGCTGCTGACTACAGGCTCGCTAGTGACGCTTCTAATGTTAAAGGG 2497  
Qy 746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765  
Db 2498 GCTATGGGCTCCAGGCTGCTGTAGTTACCCAGAAATGTAGCAAACTTGGCAGCAAT 2557  
Qy 766 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 785  
Db 2558 ATCTTATCTCTAGTCTATGCTGGCAGCTTACACTTTTACTTTGTAGTAATTTACCACTC 2617  
Qy 786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAenLysAspLys 805  
Db 2618 ATGTCTTGGTGGAAATTAATGAAATGAAACTGTTGCTGTGGTCAAGCCTTGAAGGACAAG 2677  
Qy 806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAenIleArgThrIle 825  
Db 2678 AAAGACTAGATCTCTGGGAAGATCGCTACAGAAGCAATTTGAAACTTCCGCACTGT 2737  
Qy 826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845  
Db 2738 GTCTCTTTGACTAGGAGCAGCAAGTTTGAACATATGTATGCCCAGAGCTTGCAGATACCA 2797  
Qy 846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865  
Db 2798 TACAGAAATGCTTTGAAGAAAGCAGACGCTTTTGGGATCACCTTCGCTTACCCAGGCC 2857  
Qy 866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 885  
Db 2858 ATGATTTATTTTCTATGCTGCTTGTTCGTTCCGTTCCGTTCCGTTCTTGTGGCAGCA 2917  
Qy 886 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 905  
Db 2918 CTCATGACGCTTGAATAATGTTATGTTGTTATTTCTGCTGTTGTTCTTGTGTCATGGCA 2977  
Qy 906 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 925  
Db 2978 GCAGGGAATACCAGTTTCATTCGCTCTCTAGTACGCGAAGGCCAAAGTCTCAGCATCCCC 3037

QY 926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgGlnGluGlyLysLys 945  
DB 3038 ATCATCAGGATCATGAGAAATCCCGAGATTGACAGCTACAGCAGGAGGCTTGAAG 3097  
QY 946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965  
DB 3098 CCTAATGGTTAGAGGAATGTGAATTTAATGGAGTCATGTTCAACTATCCACCCGA 3157  
QY 966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985  
DB 3158 CCCAACATCCAGTGTCTCAGGAGCTGAGTTCGAGGTGAAGAGGGGCAACGCTTCGC 3217  
QY 986 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 1005  
DB 3218 CTGGTGGCAGCAGTGGCTCGGAGAGAGTACAGTGGTCCAGTGTCTCAGCGCTTCTAC 3277  
QY 1006 AspProValGlnGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025  
DB 3278 AACCCATGGCTGGACAGTGTCTTAGATGGCAAGAAATAAACAACCTCAACGCTCCAG 3337  
QY 1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045  
DB 3338 TGCTCCGCCGCCACCTGGGCATGTGTGCCAGGAGGCCATCTCTGTGACTCGAGCATC 3397  
QY 1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065  
DB 3398 ACCGAGNACATCCCTACGAGAGCAACAGCCGCTGTCTCTCATGAGGAGATCGTGAGG 3457  
QY 1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085  
DB 3458 GCCGCCAGGAGGCCAACATCCACCTGCTCATCGACTCACTGCTGAGAAATACAACACC 3517  
QY 1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyClyGlnLysGlnArgLeuAlaIleAla 1105  
DB 3518 AGAGTGGGACAAAGGAGTCACTGTCTCGGGCGGAGAGCGCATCGCCATCGCGC 3577  
QY 1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125  
DB 3578 CGGCCCTCTGTACAGACCTCACTCACTTACTTCTGTGATGAGGACATCAGCTCTGGAT 3637  
QY 1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgCys 1145  
DB 3638 ACGGAGATGAAAGGTCTGCCAGGAGCGCTGCAGAACCCAGGAGGCCCGCATCTGC 3697  
QY 1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165  
DB 3698 ATTGTGATCGGCACCGCTGTCCACATCCAGAACGACACTTGATCGGTGATTCAG 3757  
QY 1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185  
DB 3758 AAGGCCAGCTCAAGGAGCAGCGGCCACCACACGAGCTGTCTGCCCAAGAGCATCTAT 3817  
QY 1186 PheLysLeuValAsnAla 1191  
DB 3818 TTCTCGATGGTTGAGGCT 3835

## RESULT 11

US-08-583-276-18

Sequence 18, Application US/08583276

Patent No. 5837536

GENERAL INFORMATION:

APPLICANT: McDonagh, Kevin T.

APPLICANT: Nienhuis, Arthur

APPLICANT: Tolstoshev, Paul

TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN

TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED

TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carella, Byrne, Bain, Gilfillan,

ADDRESSEE: Cecchi &amp; Stewart

STREET: 6 Becker Farm Road

CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DM4.V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,276  
FILING DATE: 05-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/332,444  
FILING DATE: 31-OCT-1994  
APPLICATION NUMBER: 07/887,712  
FILING DATE: 22-MAY-1992  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4669 bases  
TYPE: nucleic acid  
STRANDEDNESS: singular  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: Genomic DNA  
US-08-583-276-18

## Alignment Scores:

Pred. No.: 0 Length: 4669  
Score: 3497.00 Matches: 673  
Percent Similarity: 74.96% Conservative: 246  
Best Local Similarity: 54.89% Mismatches: 271  
Query Match: 58.17% Indels: 36  
DB: 2 Gaps: 9

US-09-873-409-6 (1-1195) x US-08-583-276-18 (1-4669)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValAlaCysLeuProLeuMetProLeu 20  
DB 575 ATGTGGTGGGAACCTTGGCTGCCATCATCGGGCTGGAGCTTCTCTCATGATGCTG 634  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 635 GTGTTTGGAGAAATGACAGATATCTTTGCAATGCAGGAAATTTAGACATCTGATGTCA 694  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 695 AACATCACTAATAGAGATGATATCATGAGGCTTCTTCATGAATCTGGAGGAGAC 754  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 755 ATGACACAGGTATGCTTATTATTACAGTGAATTTGGTGGGTGCTGGTGTCTGTCTTAC 814  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 815 ATTCAAGTTTCAATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
DB 875 TTTTTCATGCTATATATGCGCAGGAGATAGGCTGCTGTTGATGTGCACGATGTGGGCG 934  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 935 CTTAACACCCGACTTACAGATGATGTCTTAGATTAATGAATTTATTTGTCGCAAAAT 994  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACAG 1054  
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 1055 GGTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTGCTGCT 1114

Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	2189	TTGTCTACAGTTTCGTAATGTCAGCGTCATCGCTGGTTTCGATGAGTGGAGTCTGTCGAG	2248
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
Db	2249	AAAGGAANTCATGATGAACATCATGAAGAAGAAGGCAATTACTTCAAACCTTGTCCACAATG	2308
Qy	561	Gln-----AspIleLysAlaAspGluGlnMetGluSerMetThr	574
Db	2309	CAGACACAGAAATCAAGTTGAATTAGAAATGCAGCTGATGAATCCAAAAGTGAAATT	2368
Qy	575	TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal-----	589
Db	2369	GATCCCTTGGAAATGCTTCAAATGATTCAGATCACGCTCAATTAAGAAAAAGATCAACT	2428
Qy	590	--LysSerIleLys--SerAspPheIleAspLysAlaGluSerThrGlnSerLys	607
Db	2429	CGTAGGAGTTCGCTGGATCAACACCCCAAGACAGAAAGCTTAGTATCCAAAGAGGCTCTG	2488
Qy	608	GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTyr	627
Db	2489	GATGAAGATATACCTCCAGTTTCCTTTTGAGGAGTATGAAGCTAAATTTAACTGAATGG	2548
Qy	628	ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe	647
Db	2549	CCTTATTTCCTTGGTGTATTTTGTGCCATTATAAATGGAGGCTCCAAACCACGACATTT	2608
Qy	648	SerIleIlePheAlaLysIleIleThrMetPheGlyAsn--AsnAspLysThrThrLeu	666
Db	2609	GCAATAATATTTTCAAGATATTAGGGGTTTTTCAAGAAATGTATGATCTCTGAAACAAAA	2668
Qy	667	LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal	686
Db	2669	CGACAGAATACTAACTGTGTTTTCACATATGTTTTCAGCCCTTGGAAATTAATTCCTTTATT	2728
Qy	687	SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu	706
Db	2729	ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAAGCTGGAGAGATCCTCAACCAAGCGGCTC	2788
Qy	707	ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTyrPheAspGluLysGlu	726
Db	2789	CGATACATGTTTTTCGATCCATGCTCAGACAGAGATGTAGTGGTTGTGATGACCCATAAA	2848
Qy	727	AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla	746
Db	2849	AACACCACTGGAGCATTGACTACCGGCTCGCAATGATGCTGCTCAAGTTAAAGGGCT	2908
Qy	747	ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle	766
Db	2909	ATAGGTTCCAGGCTTGTGTGAATTATCCCAAGATATAGCAATTTCTTGGGACAGGAATAATT	2968
Qy	767	IleSerPheIleTyrGlyTyrGluMetThrPheLeuIleLeuSerIleAlaProValLeu	786
Db	2969	ATATCCTTCATCATGTTGGCAACTTAACACTGTACTCTTAGCAATTTGTACCCATCAAT	3028
Qy	787	AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln	806
Db	3029	GCAATACAGAGTGTGTGAATGAATAATGTTGTCTGGACAGCACTGAAAGATATAGAAA	3088
Qy	807	GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal	826
Db	3089	GAAC TAGAGGTCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAAACCGTTGTT	3148
Qy	827	SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis	846
Db	3149	TCTTTGACTCAGGACGAGAAGTTTGAACATAATGTTATGTCTCAGAGTTTTCACGGTACCATAC	3208
Qy	847	ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe	866
Db	3209	AGAAACCTCTTTGGAAGACACACATCTTTGGAATTACATTTTCTCTCACCAGGCAATG	3268
Qy	867	IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg	886

```

:::|||||:::|||||
3269 ATGTAATTTTCCTATGCTGGATGTTCCGGTTGGAGCCTACTTGGTGGCACATAACTA 3328
QY MetThrProGluGlyMetPheIleValPheThrAlaIleAlaIleAlaIle 906
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTCAGCTGTTGCTTTGGTGGCATGGCCGG 3388
QY GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3389 GGGCAAGTGAAGTTCATTGCTCTGACTATGCCAAGCCCAATATATCAGCAGCCACATC 3448
QY PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3449 ATCATGATCATTTGAAAAACCCCTTTGATTGACAGTACAGCAGGAAGCCCTAATGCCG 3508
QY AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3509 AACACATTGGAAGGAATGFCACATTTGGTGAAGTTGTTATTCAACTATCCACCCGACCG 3568
QY AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3569 GACATCCAGTCTTCAGGAGCTGAGCCTGGAGGTGAGAGGCGCCAGAGCTGGCTCTG 3628
QY ValGlySerSerGlyCysGlyLysSerThrSerValIcInLeuLeuGlnArgLeuTyrAsp 1006
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3629 GTGGCAGCAGTGGCTGTGGGAAGACACAGTGTCCAGCTCTCGAGCGGTTCTACGAC 3688
QY ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTyr 1026
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3689 CCCTTGGCAGGGAAGTGGCTGTGATGGCAAGATAAAGCGACTGAATGTCAGTGG 3748
QY LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3749 CTCGAGCACACCTGGCATCTGTCAGAGAGCCATCTCTGTTGACTGCACATTTGCT 3808
QY GluAsnIleAlaTyrGlyAspAsnSerArgValProLeuAspGluIleLysGluAla 1066
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3809 GAGAACATTGCTATGGACACACAGCCGGTGTGTACAGGAGAGATCGTGGAGGCA 3868
QY AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3869 GCAAGAGAGCCCAACATACATGCTCTTCATCGAGTCACTGCTTAATAATATAGCACTAA 3928
QY ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db GTAGGACAAAGGAAGTCTGCTGTTGGCCAGAAACACAGCATTTGCCATAGCTCGT 3988
QY AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
3989 GCCCTTTAGACAGCCTCATATTTTCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048
QY AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db GAAAGTGAAAGGTTGTCCAAGAGCCCTGGCAAAACCCAGAAAGCCGCCCTGCAAT 4108
QY ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4109 GTGATTCCTCACCCCTGTCCACCATCCAGATGCACACTTAATAGTGGTTCAGAAAT 4168
QY GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4169 GGCAGAGTCAAGCAGCAGTGGCAGCATCAGCAGCTGTGGCACAGAAAGGCATCTATTT 4228
QY LysLeuValAsnAlaGln 1192
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
4229 TCAATGCTCAGTGTCCAG 4246
```

## RESULT 12

US-08-996-545-1

; Sequence 1, Application US/08996545

; Patent No. 5928898

; GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.

```

; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at rD of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
; US-08-996-545-1

```

## Alignment Scores:

Pred. No.:	7,23e-247	Length:	4002
Score:	2282.50	Matches:	494
Percent Similarity:	59.70%	Conservative:	232
Best Local Similarity:	40.62%	Mismatches:	446
Query Match:	37.97%	Indels:	45
DB:	2	Gaps:	11

US-09-873-409-6 (1-1195) x US-08-996-545-1 (1-4002)

QY	15	LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeu	34
Db	368	GTGCCATTGCTGCGGCTCGACTTCCAGA-----GGATAATGTTAT	409
QY	35	ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrTyrVal	50
Db	410	ATCAAAATCTCGTACGAC-GAGTTCTATGATGATTCACCAAGAACCTACTGTACTTCGTA	468
QY	51	GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTyrIleThr	70
Db	469	TACTCGGTATCGGCAGTTTGTCTACTGTCTATGTTAGTACTGTTGGTCTCACTATACC	528
QY	71	AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheIleSerValLeuAlaGlnAsp	90
Db	529	GGAGAACACGCCACGCAAGATATTCCTTGAGTCTATCTCTGCGCCAGAAC	588
QY	91	IleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle	109
Db	589	ATTGGCTATTTTGTATAAACTCGGTGCGGGGAAGTGACCCCGGTATACAGCCGATACA	648

QY 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaIleLeuPheGlnAsnMetSerThr 129  
DB 649 AACCTTATCCAGATGCGCATTTCCGAGAAAGGTCTCTCATTTGTAGCTCCCTGGCGACA 708  
QY 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrIlePheLeuThrLeuValThrLeu 149  
DB 709 TTCGTACAGCATTCAATATCGCTACGTCATATACCTGGAAGTGTGCTCTTAATTTCCAGC 768  
QY 150 SerThrSerProLeuMetAlaSerAlaAlaAlaCysSerMetValIleSerLeu 169  
DB 769 TCAACAATTTGGCCCTCGTTCTCACCATCGGGCGGTCTCTCAGTTTATCATCAAGTAC 828  
QY 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189  
DB 829 AGCAAAAGTCGCTTCACAGCTACCGTGCAGCGCGCACTGTTGCGGAAGAGGTTCATCAGC 888  
QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209  
DB 889 TCCATCAGAAATGCCACAGGTTTGGCACCACAGACAGCTTCGGAAGCAGTATGAGGTC 948  
QY 210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229  
DB 949 CACTTAGCAGAGCTGAGAAATGGGAAACAAAGAACCCAGATTGTCATGGTTTCATGATT 1008  
QY 230 GlyAlaValTyrPheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSer 249  
DB 1009 GGGCCATGTTTTGGCCTTATGACTCGCAATACGGTCTTGGCTTCCTGGATGGGTTCTCGT 1068  
QY 250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269  
DB 1069 TTCCTGGTAGAGT-----GCAGTCGATGTGGGTGATATCTCAGTTCTCATGGCC 1122  
QY 270 ValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289  
DB 1123 ATCTTGATCGGATGTTCTCTTGGGGAAGTTAGTCCAAATGCTCAAGCATTTTCAAAAC 1182  
QY 290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309  
DB 1183 GCTGTGGCGCGCGCAAGATATTTGGAACGATCGATCGCCATTCAGTCCATTTAGATCCA 1242  
QY 310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329  
DB 1243 TATTCGAACGAAGGAAGACGCTCGACCATTTTGGGGCCACATGAGTTACGCAATGC 1302  
QY 330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349  
DB 1303 AAGCATATTTACCATCTAGACCCGAGGTCCCGTCATGAGGATGTTCTCTGTCAATG 1362  
QY 350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369  
DB 1363 CCGCTGGAAAAACAACCGCTTTAGTCGGCCCTCTGGCTCTGGAAGAAAGTACGGTGGTC 1422  
QY 370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389  
DB 1423 GGTCTGGTTGAGCGATTCTACATGCTCTGCGGTACCGGTTTGTGATGGCCATGAC 1482  
QY 390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409  
DB 1483 ATCAAGACCTCAATCTCCGTGGCTTCGCAACACATCTCTTTGGTTAGCCAGGAGCCT 1542  
QY 410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly----- 423  
DB 1543 GTTCTTTTGGCAGCAGCATTTAAGATATTAGGACGGTCTCATCGGCACAAAGTAC 1602  
QY 424 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
DB 1603 GAGAATGAATCCGAGATAAGTCCGGAACTCATCGAAGCCGCAAAATGGCGAAT 1662  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
DB 1663 GCTCATGACTTATTAATGCTTCCTGAGGTTATGAGCAATGTTGGGAGCGGTGGC 1722  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480

DB 1723 TTTTCTCTTTAGGTGGCCAGAAACAGCGCATTCGAATCGCCGTCCTGTTAGTGAC 1782  
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
DB 1783 CCAAAAATCTCTCTCTGATGAAGTACTTCCGGCTTGGACACAAATCCGAAGCGGTG 1842  
QY 501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520  
DB 1843 GTTCAAGCAGCTTTGAGAGGCGCAGCTGAAGCGCAACTTATTTGTGATCGCTCATCGC 1902  
QY 521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540  
DB 1903 CTTTCCAGCATCAAAACGGCGCACAACTTGTGTTCTGTCATAGGCAAAATTTGTGAA 1962  
QY 541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560  
DB 1963 CAAGNACTCAGATGAATTTGTTGACCGCGGAGCGCTTATCGCAAACTTGTGGAGGCT 2022  
QY 561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580  
DB 2023 CAACGTATCAATGAAACAGAAAGGAAGCTGACGCTTGGAGGACGCCGCTGAGGATCTC 2082  
QY 581 ThrAsnSer-----LeuProLeuHisSerValLysSerIleLysSerAspPhe----- 596  
DB 2083 ACGAATGCGAGATATTGCCAAATCAAACTGCGTCAAGCGCATCATCCGATCTCGACGGA 2142  
QY 597 -----IleAspLys-----AlaGluGluSerThrGlnSerLysGluIleSer 610  
DB 2143 AAACCCACAACCATTTGACCGCAGCGGACCCCAAGCTCTGTTTCCAGCGCGATCTTTCT 2202  
QY 611 -----LeuProGluValSerLeuLeuLysIleLeuLys----- 621  
DB 2203 AAAAGACCCCGCAAAACAACTCCGAAATACTCATTTATGACGCTGCTCAAAATTTGTGCT 2262  
QY 622 ---LeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640  
DB 2263 TCTTCAACCGGCTGAAATCCCGTACATGCTCATCGGTCCTTGTCTCTCAGTGTAGCT 2322  
QY 641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly--- 659  
DB 2323 GTTGTGGCCAAACCCAGCAAGCATGCTATATGTAAAGCCATCAGCACATCTCGCTC 2382  
QY 660 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 678  
DB 2383 CCAGAATCAACAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTTCATTGATGTTCT 2442  
QY 679 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 698  
DB 2443 GTGGTTGGTATCATTCAGTTTATCAGCAGTCAACCAATGGTGTGTCATTTGCCGTATGC 2502  
QY 699 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 718  
DB 2503 TCGAGAGACTTATTCGTCGCGGAGAACACTGCTTTCGGACGATACTCCGTCAAGAC 2562  
QY 719 IleAlaTyrPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 738  
DB 2563 ATTGCTTTCTTTGACAAGGAAGAATAGCACCGGCGCTCTGACCTCTTCTGTCCACC 2622  
QY 739 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758  
DB 2623 GAGACGAACATCTCTCCCGGTGTAGCGGTGACTCTAGGCAAGCATCTTGATGACCTCC 2682  
QY 759 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTyrGluMetThrPheLeu 778  
DB 2683 ACGACCTAGGACGCTATCATTTATTCCTGGCGATTTGGGTGGAAATTTGGCTTAGTT 2742  
QY 779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798  
DB 2743 TGTATCTCGTTGTGCGGTTCTCTCGTGGCATCGGTTTCTACCGATTTCTATATGTCAGCC 2802  
QY 799 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 818

Db 2803 CAGTTTCAATCAACGCTCCAAAGCTTGTATGAGGGATGTCAAACTTTGCTTCGAGGCT 2862  
Qy 819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838  
Db 2863 ACATCGTCTATCCGACAGTGTGCTTAATACCCGGGAAGGATGCTCTGGAGATTTC 2922  
Qy 839 GluGluMetLeuGlnThrHisArgAsnThrSerTyrLysAlaGlnIleLeuGlySer 858  
Db 2923 CATGCCAGCTTGACGCACACAGCAGGACGAGCTAATCTCTCTGTGAGGTATCCCTG 2982  
Qy 859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878  
Db 2983 TTATATCGTCTGCGCAGGACACTGTCTTCTCTCGCTGCGCTCGGGTTTGTGTACGA 3042  
Qy 879 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898  
Db 3043 GGCACACTCTTGCTTCCACACGAGTATGACATTTTCGCGTCTTTGTGTTCTCTCCGAG 3102  
Qy 899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918  
Db 3103 ATTCTCTTGTGCTCAATCCGCGGACCGCTCTTTCTTTCACACAGACATGGGCAAG 3162  
Qy 919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 938  
Db 3163 GCGAAGAAATCGGCGCGCGAATTCGACGACCTGTTCACCGAAAGCCACAAATTGATAAC 3222  
Qy 939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958  
Db 3223 TGGTCTGAAGAGGCGAGAGACTCGAAGCGGTGAAGTGAATTCGATTTAGAACGCTG 3282  
Qy 959 SerPheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978  
Db 3283 CACTTCAGATACCGCCGCCAGACAGCTGCTCTCGCGGCTTGGACCTGACCGTG 3342  
Qy 979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998  
Db 3343 AAGCCTGGACAATATGTTGGCTTGTTCGACCCAGCGGTGTGCGCAAGATGACCACT 3402  
Qy 999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018  
Db 3403 GCATTGCTTGAGCGCTTTACGATGCGATTCGCCGGTCCATCTTGTGATGGGAAGAC 3462  
Qy 1019 AlaLysLeuLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038  
Db 3463 ATAAGTAACTAATATCACTCTACCGCAGCTTCTGTCACTGCTCAGCGAGCGC 3522  
Qy 1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058  
Db 3523 ACACGTACACAGGCGCACCATCAAGAAACATCTTACTTGTGATTTGTGAGATGACGTA 3582  
Qy 1059 ProLeuAspGluIleLysGluAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078  
Db 3583 CCGAAGAAATCTTGATTAAGCTTGGCAAGACCGCTAATATCTACGACTTCATCATGTGC 3642  
Qy 1079 LeuProGluLysTyrAsnThrGlnValGlyLysGlyAlaGlnLeuSerGlyGlyGln 1098  
Db 3643 CTCGCGGAGGCTTAAATACAGTTGTTGGCAGAGGAGGAGCATGTTGCTCGCGGCGAA 3702  
Qy 1099 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuLeuAsp 1118  
Db 3703 AAGCAAGCTGTGGCGCATTTGCCGAGCCCTTCTCGGATCCCAAAATCCTTCTCTCGAT 3762  
Qy 1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138  
Db 3763 GAACGAGCTCAGCCCTCGACTCCGAGTCAGAAAGGTCTGCCAGGCGCTTTGGATGCC 3822  
Qy 1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158  
Db 3823 GCTGCCGAGGCGGACCAATCGCGCTTGACACCGACTCAGCAGCATTCACAAAGCGC 3882  
Qy 1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178  
Db 3883 GACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCAAGCAAGCTG 3942

Qy 1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194  
Db 3943 GTCCAGAAAAGCGCGGTACTACGAGCTGTGTCACACTTCGAGAGCTTG 3990

## RESULT 13

US-08-996-545-3  
; Sequence 3, Application US/08996545  
; Patent No. 5928898  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atxD of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,545  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4002 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: mRNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
US-08-996-545-3

Alignment Scores:  
Pred. No.: 7,23e-247 Length: 4002  
Score: 2282.50 Matches: 494  
Percent Similarity: 59.70% Conservativeness: 232  
Best Local Similarity: 40.62% Mismatches: 446  
Query Match: 37.97% Indels: 45  
DB: 2 Gaps: 11

US-09-873-409-6 (1-1195) x US-08-996-545-3 (1-4002)

Qy 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34  
Db 368 GUGCAUUGCUGCGCGUCGACUUCUCCAG- - - - -GGAUAAUGUUAU 409  
Qy 35 ValGlnThrAsnThrTyrSerPhePhe- - - - -ArgLeuThrLeuTyrTyrVal 50  
Db 410 AUCAAUUGUCUGACGAC-GAGUUCUAUGAUAUAGCAAGAACGACGACUUGUA 468  
Qy 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTyrIleThr 70  
Db 469 UACCUCGGUAUCCGCGAGGUUGUACUGUACUGUANGUACUUCUUGGCUCAUCUAUACC 528







Ds 2683 ACGACCCUAGGAGCGGCUAUAUUGCCUGGCAUUGGUGGAAUUGGCCUUAUUU 2742  
Qy 779 ILeuSerIleAlaProValIleAlaValThrGlyMetIleGluThrAlaAlaMetThr 798  
Ds 2743 UGUUUCUGGUGGCGGCUUUCUGGCGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 2802  
Qy 799 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 818  
Ds 2803 CAGUUUCAAUCACGCUCAAGCUGUUAUGAGGGAUGUCAAACUUGUUGCGAGGCU 2862  
Qy 819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetThr 838  
Ds 2863 ACAUCGUCUACCGCAGUUGUUAUACCGGGAAGGAGUUGUGGGAUUAUAC 2922  
Qy 839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleLeuGlySer 858  
Ds 2923 CAUGCCCGGUGGACGACAGGACGAGGACGACUUAUCUUGGUGGUGGUGGUGGUGG 2982  
Qy 859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878  
Ds 2983 UUAUAGUGGUGGUGGACGACUUGUUAUUCUGGUGGUGGUGGUGGUGGUGGUGG 3042  
Qy 879 AlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898  
Ds 3043 GGGACACUUCUGGUGACACGACGAGUAUGACAUUUUUGGUGGUGGUGGUGGUGG 3102  
Qy 899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918  
Ds 3103 AUUCUUCUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3162  
Qy 919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuGluLysLysProAsnIleAspSer 938  
Ds 3163 GCGAGAAUUGGCGCGCGGCGGAAUUCGACGACUGUUGGACCGGACCAAAUUGAUAC 3222  
Qy 939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958  
Ds 3223 UGGUGUAGAGGCGGAGAGACUGCUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3282  
Qy 959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978  
Ds 3283 CACUUCAGAACCGGACCGCGCGGACGACGUGUUGGUGGUGGUGGUGGUGGUGGUGG 3342  
Qy 979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998  
Ds 3343 AAGCCUGGACAAUUAUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3402  
Qy 999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018  
Ds 3403 GCAUUGCUGAGCGCUUUAUGAUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3462  
Qy 1019 AlaLysGluLeuAsnValGlnTyrLeuArgSerGlnIleAlaIleValProGlnGluPro 1038  
Ds 3463 AUAAGUAAACUAAUUAUACUUCUACCGGACGUGUUGGUGGUGGUGGUGGUGGUGG 3522  
Qy 1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058  
Ds 3523 ACACUGUACGAGGCGACCAUACAGGAAACAUUUAUUGGUGGUGGUGGUGGUGGUGG 3582  
Qy 1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078  
Ds 3583 CCGGAAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3642  
Qy 1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGln 1098  
Ds 3643 CUCCGAGGCGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3702  
Qy 1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118  
Ds 3703 AAGCAACUGUGGCGCAUUGCCGAGCCUUCUUGGAGUCCCAAAUUCUUCUUCGAG 3762  
Qy 1119 GluAlaThrSerAlaLeuAspAsnAspSerGlnLysValValGlnHisAlaLeuAspLys 1138  
Ds 3763 GAAGCGACGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3822

Qy 1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158  
Ds 3823 GCUGCCCGAGGCGGACCAACCAUAUUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3882  
Qy 1159 AspLeuIleValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178  
Ds 3883 GAGUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3942  
Qy 1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194  
Ds 3943 GUCCAGAAAAGGCGGUGUACGAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 3990

## RESULT 14

US-09-328-320-1  
; Sequence 1, Application US/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/328,320  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4002 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..4002  
US-09-328-320-1

Alignment Scores:  
Pred. No.: 7.23e-247 Length: 4002  
Score: 2282.50 Matches: 494  
Percent Similarity: 59.70% Conservative: 232  
Best Local Similarity: 40.62% Mismatches: 446  
Query Match: 37.97% Indels: 45  
DB: 4 Gaps: 11

Qy	370	GlnLeuLeuGlnhrgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp	389
Db	1423	GGCTTGGTTGAGCGAATCTTACATGCTGTTCGCGTAGCGGTTTTCGTGGATGGCCATGCAC	1482
Qy	390	IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro	409
Db	1483	ATCAGGACCTCAATCTCCGCTGGCTTCCCAACAGATCTCTTTGGTTAGCCAGGAGCCT	1542
Qy	410	ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly	423
Db	1543	GTTCCTTTTGGCAGCAGATTTATAGAATAATTAGGCACGGTCTCATCGGCACAAAGTAC	1602
Qy	424	-----ArgAspAspValThrAspGluGluMetGluuArgAlaAalaArgGluuAsn	440
Db	1603	GAGAAATGAATCCGAGGATAAGTCCGGGAATCATCTCAGAAACGGCGCAAAAATCGCGAAT	1662
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1663	GCTCATGACTTTATTACTGCTTGCCTGAGGTTATGAGACCAATGTTGGCAGCGTGGC	1722
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1723	TTTTCTCTTTCAGTGCCAGAAACAGCGCAATTCGAATCGCCGTCGCGCTTTGTTAGTGAC	1782
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1783	CCAAAAATCTGCTCTCGATGAAGTACTTTCGGCTTGGACACAAAAATCCGAGGGCGTG	1842
Qy	501	ValGlnAlaAlaLeuGlnLysAlaSerLysGlyArgThrThrIleValValAlaHisArg	520
Db	1843	GTTCCAACGACGTTTGAGAGGGCAGCTGAAGCGCCGAATCTATTGTGATCGCTCATCGC	1902
Qy	521	LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu	540
Db	1903	CTTTCCACGATCAAAAGCGGCACACATTTGCTGTCTGCTCAATGGCAAAATGCTGAA	1962
Qy	541	LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer	560
Db	1963	CAAGGAATCACGATGAATTTGGTTGACCGCGAGGCGCTTATCGCAAACTTTGTGGAGCT	2022
Qy	561	GlnAspIleLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys	580
Db	2023	CAACGTATCAATGAACAGAAAGGAGCTGACGCTTGGAGGACCGCGCTGAGGATCTC	2082
Qy	581	ThrAsnSer-----LeuProLeuHisSerValLysSerIleLysSerAspPhe	596
Db	2083	ACGAATCGGATATTCGCAAAATCAAAATCTGGCTCAAGCGCATCATCCGATCTCGACGA	2142
Qy	597	-----IleAspLys-----AlaGluGluSerThrGlnSerLysGluIleSer	610
Db	2143	AAACCCACACCATTCACCGCACGGGCACCCACAAGTCTGTTTCCAGCGCATTTCTTCT	2202
Qy	611	-----LeuProGluValSerLeuLeuLysIleLeuLys	621
Db	2203	AAAAAGACCCCGCAAAACAACTCCGAAATAACTCATTTATGGACGCTCTCTCAAAATTTGTGCT	2262
Qy	622	---LeuAsnLysProGluTyrProPheValValLeuGlyThrLeuAlaSerValLeuAsn	640
Db	2263	TCCCTTCAACCGCCCTGAATTCGCTACATGCTCATCGGCTTGTGCTTCTTCAGTGTTACT	2322
Qy	641	GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly	659
Db	2323	GGTGGTGGCCAAACCCAGCAAGCAGTGTATATGTCTTAAGCATCAGCACACTCTCGCTC	2382
Qy	660	---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal	678
Db	2383	CCAGAATCACAATATACGACGCTTCGCACATGATGCGAATTTCTGTGCTCATTTGATGTTCTC	2442
Qy	679	IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla	698
Db	2443	GTGGTGGTATCATTCAGTTTATTCATTCACGCGAGTCAACCAATGCTGCTGCATTTGCCGTATGC	2502

QY 699 GlyGluLeuLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 718  
DB 2503 TCCGAGAGACTATTTCGTCGGCGAGACACTGCCTTTTCGGACGATATCCGTCGAAGC 2562  
QY 719 IleAlaTrpPheAspGluLysGlnAsnSerThrGlyLeuThrThrIleLeuAlaIle 738  
DB 2563 ATTGCTTTCTTACAGGAAGAGATAGACCGCGCTCTGACCTCTTTCCTGTCCACC 2622  
QY 739 AspileAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758  
DB 2623 GAGACGAGCATCTCCGCTGTAGCGGTGACTCTAGCGAGATCTGTAGCACTCC 2682  
QY 759 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 778  
DB 2683 ACAGCCCTAGGAGCGGTATCATATTATGCGCTGCGGATTCGGTGGAAATGCGCTTAGTT 2742  
QY 779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798  
DB 2743 TGTATCTCGGTTGCGGTTCTCTGCGATGCGGTTCTTACCGATTTCTATATGCTAGCC 2802  
QY 799 GlyPheAlaAsnLysAspLysGlnLeuLysHisAlaGlyLysIleAlaThrGluAla 818  
DB 2803 CAGTTTCATACGCTCCAGACTGCTTATGAGGATCTCAAACTTGTCTCGAGGCT 2862  
QY 819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838  
DB 2863 ACATCGTCTATCCGACAGTTGCGTCAATTAACCGGGAAAGGATGCTCGGAGATTTAC 2922  
QY 839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858  
DB 2923 CATGCCAGCTTGACGACAGGACGAGCAGTCTAATCTCTGTGTGAGGTATCCCTG 2982  
QY 859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878  
DB 2983 TTATGCGTCTGCGCGAGCACTTGTTTCTTCTGCGTCCGCTCGGTTTGTGACGA 3042  
QY 879 AlaTyrIleuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898  
DB 3043 GGGACACTTCTGTCACACGAGTATGACATTTTCGCTCTTGTGTTTCTCGAG 3102  
QY 899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918  
DB 3103 ATTCTCTTTGCTGCTCAATCCGCGGCGACCGTCTTTCTCTTCCACGACATGGGCAAG 3162  
QY 919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLysLysProAsnIleAspSer 938  
DB 3163 GCGAAGATCCGCGCGCGAATTCGAGACTGTTCCAGCAAGCCAAATGTATAAC 3222  
QY 939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958  
DB 3223 TGGTCTGAAGAGGCGGAGAGCTCGAAACGCTGGAAGTCAATCGAATTTAGGAACGTG 3282  
QY 959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerIle 978  
DB 3283 CACTTCAGATACCGACCGCCCGAGAACAGCTCTCTGCGCGCTTGGACCTGACCGGTG 3342  
QY 979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998  
DB 3343 AAGCTTGACAAATATGTGCGTTGCGGACCGCGGTGTGCGAAGATACACCAAT 3402  
QY 999 GlnLeuGlnArgLeuTyrAspProValGlnGlnValLeuPheAspGlyValAsp 1018  
DB 3403 GCATTCCTGAGCGCTTTTACGATGCGATTGCGGGTCCATCTCTGTGTGATGGAAGGAC 3462  
QY 1019 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038  
DB 3463 ATAAGTAACATAATATCAACTCTCTCCGCGAGCTTTCTGTCACTGTGTCAAGGAGCGG 3522  
QY 1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058  
DB 3523 ACATGTACCGAGGCGCATCAAGAGAAACATCTTACTTGTATGTGCAAGATGACGTA 3582  
QY 1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078

DB 3583 CCGAAGAAATCTTGTATTAGGCTTGCAGAGCGCTTAATATCTACGACTTCATATGTCG 3642  
QY 1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1098  
DB 3643 CTCGCGAGGCGCTTTATACAGTTGTGGCAGCAAGGAGGATGTTGTCTGCGGCCAA 3702  
QY 1099 LysGlnArgLeuAlaIleAlaArgAlaLeuGlnLysProLysIleLeuLeuAsp 1118  
DB 3703 AAGCAACGTTGGCCATTGCCGAGCGCTTCTTCGGGATCCCAAAATCTTCTTCTCGAT 3762  
QY 1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138  
DB 3763 GAAGCGACGCTCAGCCCTCAGCTCCGAGTCAGAAAGGTCTCCAGCGGCTTTGGATGCC 3822  
QY 1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158  
DB 3823 GCTGCCGAGGCGGACCAACAATCGCGTTGCACCGGACTCAGACGATTCAAAAGGCG 3882  
QY 1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178  
DB 3883 GAGTTATCTATGTTTTCGACCAAGCGCAAGATCGTCGAAAGCGGAACGACACGCACTG 3942  
QY 1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194  
DB 3943 GTCCAGAAAAGGCGCGGTACTACGAGCTGGTCAACTTGCAGAGCTTG 3990  
RESULT 15  
US-09-328-320-3  
; Sequence 3, Application US/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atzD of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/328,320  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4002 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: mRNA  
; HYPOTHETICAL: NO



[illegible][illegible]

Search completed: March 31, 2003, 15:36:39  
Job time : 436.002 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 196.92 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739  
Sequence: 1 MILGILASLVNGACPLMPL.....SEKSAVOAALKDTPRYSF 541

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL-frame+ p2n.model -DEV=xlh

-Q/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003.115424\_19290/app\_query.fasta\_1.7544  
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09873409 @CGN 1.1 960 @runat\_27032003.115424.19290  
-NCPU=6 -ICPU=3 -NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2737	99.9	2021	10	US-09-873-409-15
2	2697	98.5	3702	10	US-09-873-409-13
3	2568.5	93.8	1940	10	US-09-873-409-16
4	2528.5	92.3	3621	10	US-09-873-409-14

5	1868	68.2	3177	10	US-09-873-409-12	Sequence 12, Appl
6	1566.5	57.2	4189	10	US-09-866-866A-5	Sequence 5, Appl
7	1555.5	56.8	3860	10	US-09-866-866A-1	Sequence 1, Appl
8	1552.5	56.7	4317	9	US-10-044-671-1	Sequence 1, Appl
9	1548	56.5	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
10	1546.5	56.5	3860	10	US-09-866-866A-3	Sequence 3, Appl
11	1546.5	56.5	8630	10	US-09-306-417-1	Sequence 1, Appl
12	1546.5	56.5	8630	10	US-09-306-417-2	Sequence 2, Appl
13	1543	56.3	4425	10	US-09-769-097-3	Sequence 3, Appl
14	1541	56.3	4369	10	US-09-769-097-1	Sequence 1, Appl
15	1539	56.2	4643	9	US-10-072-621-2	Sequence 2, Appl
16	1536	56.1	4788	10	US-09-866-866A-7	Sequence 7, Appl
17	1455.5	53.1	3924	10	US-09-880-107-2299	Sequence 2299, Ap
18	1421.5	51.9	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	1220	44.5	5010	10	US-09-917-800A-483	Sequence 483, App
20	1141.5	41.7	1175	10	US-09-873-409-11	Sequence 11, Appl
21	1101.5	40.2	2856	10	US-09-873-409-10	Sequence 10, Appl
22	1028	37.5	1810	10	US-09-749-340-4	Sequence 4, Appl
23	1028	37.5	4175	10	US-09-749-340-3	Sequence 3, Appl
24	993.5	36.3	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
25	942	34.4	3861	9	US-09-938-842A-263	Sequence 263, App
26	914	33.4	4002	9	US-09-758-828-1	Sequence 1, Appl
27	914	33.4	4002	9	US-09-758-828-3	Sequence 3, Appl
28	891.5	32.5	4653	9	US-10-101-388-2	Sequence 2, Appl
29	812.5	29.7	2066	10	US-09-873-409-9	Sequence 9, Appl
30	785.5	28.7	3792	9	US-09-882-694-10	Sequence 10, Appl
31	754	27.5	2681	10	US-09-749-340-8	Sequence 8, Appl
32	754	27.5	8036	9	US-10-101-388-1	Sequence 1, Appl
33	743	27.1	3512	10	US-09-749-340-7	Sequence 7, Appl
34	738	26.9	2698	10	US-09-749-340-5	Sequence 5, Appl
35	688.5	25.1	3999	9	US-09-882-694-9	Sequence 9, Appl
36	657.5	24.0	2298	9	US-10-156-239-18	Sequence 18, Appl
37	657.5	24.0	2298	9	US-09-795-693-18	Sequence 18, Appl
38	657.5	24.0	3408	10	US-10-156-239-16	Sequence 16, Appl
39	657.5	24.0	3408	10	US-09-795-693-16	Sequence 16, Appl
40	657.5	24.0	3512	9	US-10-072-621-1	Sequence 1, Appl
41	596	21.8	2674	10	US-09-917-800A-1500	Sequence 1500, Ap
42	586.5	21.4	1749	10	US-09-815-242-9908	Sequence 9908, Ap
43	582.5	21.3	1749	10	US-09-815-242-6018	Sequence 6018, Ap
44	579.5	21.2	2247	9	US-09-870-759-40	Sequence 40, Appl
45	574.5	21.0	1737	10	US-09-815-242-8403	Sequence 8403, Ap

ALIGNMENTS

RESULT 1  
US-09-873-409-15  
; Sequence 15, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 2021  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (723)..(723)  
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)  
US-09-873-409-15

Alignment Scores:  
Pred. No.: 3.3e-288 Length: 2021  
Score: 2737.00 Matches: 541  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%  
Query Match: 99.93%  
DB: 10  
Indels: 0  
Gaps: 0  
Mismatches: 0

US-09-873-409-7 (1-541) x US-09-873-409-15 (1-2021)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCCTGGGTATACCTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAATCAACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaLeuLeuPheGly 60
DB 154 TCCTTCTCAGGTGACCTGTATTATGTTGGATAGTGTGCTGCTTGATTTTGGT 213
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGGATTATTAACCTGCAGCAGCAGACCAAGAGGATTCGAAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACCTTAACACTCGCATGACAGACATTTGACAAATCAGTATGATGATGAGATTAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTTGTTCAAAACATGCTACTTTTTTCGATTGGCCTGGCAGTTGGTTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGCTGGAAACTCACCTTAGTGACTCTATCACAGTCTCCTCTTAATAATGGCTTCAGCGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATTTGACCAAGTAAGGAATTAAGTGCTCTATTCCAAAGCT 573
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGGCTGTGCAGAGAAGCTTGTGTCATCAATCCGAACAGTCATAGCTTTAGGGCCAC 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 634 GAGAAAGAACTTCAAGGTCTTTTCTTTTAAATATAACAAGATATGCTGCTTTTATTTT 693
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 694 CCCAGTGGCTACTAAGTTGGTTCTGTTTGTAGGATATACACAGAATCTCAAGAT 753
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 754 GCAAGGATTTGGCATAAAAAGGACATAGCTTCAAGAGTCTCTTGGTGTGTGTAC 813
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 814 TTCTTTATGAATGGAACCTATGGAATGCTGCTTTTGGTATGGAACCTCTCTGATTTCTTAAT 873
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 874 GGAGAACTCGATATACCATCGGACCTGTTCTGCTGTTTCTTTAGTGAATCCATAGC 933
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 934 AGTATTGCAATGGAGCAGCAGTCCCTCACTTGAACCTTCGAATAGCCCGGAGGCT 993
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 994 GCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053
```

```
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1054 GGATATAAACCTTGAATCCATAGAGGAACCTGTGGAATTTAAATAATGTTCTTCAATTAT 1113
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAAATCTCAGAAATTAAGCTCGGAG 1173
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
DB 1174 ACAGTCGCTTGTGCTGCTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1234 AGCTTATATGATCCGATGATGCTGCTTATCATGCTGATGAGATGCATCAGAGCTTTA 1293
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1294 AATGTGCGGCATTTATCGAGACCATATTTGGAGTGTGTTAGTCAAGAGCCTGTTTGTTCGG 1353
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
DB 1354 ACCACCATCAGTAACCAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAG 1413
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1414 AGAGCAGCAGGAGGAGCAATGCTATGATTTTATCATGGAGTTTCTTAATAAATTTAAT 1473
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1474 ACATTGTTGGGGAAAAAGGAGCTCAAAATGAGTGGAGGGCAGAAAACAGAGGATCGCAATT 1533
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1534 GCTCTGTCCTTAGTTCGAACCCCAAGATTTCTGATTTTAGATGAGGTAGCTGTGCCCTG 1593
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 540
DB 1594 GATTTCAGAAAGCAAGTCAGTCGTTTCAAGCTGCACTGGAGAGGATACCCCGAGTATTCA 1653
QY 541 Phe 541
DB 1654 TTT 1656
```

## RESULT 2

```
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 13
; LENGTH: 3702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.: 1,928-283 Length: 3702
Score: 2697.00 Matches: 534
Percent Similarity: 100.00% Conservative: 0
```



Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.47% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-7 (1-541) x US-09-873-409-13 (1-3702)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTCGGGTATATCGCATCACTGGTCAATGGAGCGCTGCTCTCTTTAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATAGTGATTAACCTTATTAGTGGATGCTAGTCCAAACTAACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
DB 154 TCTTTCTTCAGGTTCACCTGATATTATGTTGGAATAGGTGTTGCTGCTTGTATTTTGT 213
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
DB 214 TACATACAGATTTCTTGTGGATTATAACTGCAGCAGCAGACCAAGAGGATTCCGAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTTCATTCAGTTTTGGCAGAGGACATCGGCTGGTTGTATAGCTGTGCATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 334 GAACTTAACACTCGCATGACAGACATTTGACAAATCAGTGATGTTTGGAGATAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCATTTGGCTGGCAGTTGGTTGGTGAAG 453
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GCGTGAACACTCACCTAGTACTCTATCCAGCTCTCTTATATATGCTTCAGCGGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGTCATCTCATTTGACCAGTAAGGAATTAAAGTCCCTATCCAAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGGCTGTGGCAGAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAG 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 634 GAGAAAGAACCTTCAAAAGGCTTTTCCCTTTTAAATATAAACAAGATATGCTTGGTTTATTTT 693
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 694 CCCAGTGGCTACTAGTTGTGTCTGTTNTTGTAAAGGTATACACAGAAATCTCAAGAT 753
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 754 GCAAAAGGATTTTGGCATAAAGAGCTATAGCTTCAAAAGTGTCTCTGCTGTGTGTAC 813
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 814 TTCTTTATGAATGGAAACCTATGGACTTGCTTTTGTGTGGAACCTCTTGATTCCTTAAT 873
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 874 GGGAACCTGGATATACCATCGGAGCTGTTCTTGCTGTTTCTTTAGTGTATCCATAGC 933
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 934 AGTTATTGCAATGGAGCAGTCCCTCACTCTTGAACCTTTCGAATAGCCCGAGGAGCT 993
QY 321 AlaPheHisIlePheGlnValIleAspLysProSerIleAspAsnPheSerThrAla 340
DB 994 GCCTTTTCATATTTCCAGGTTATTGTAGAAACCCAGTAGATAGATAACTTTTCCAGCT 1053
```

```
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1054 GGAATATAAAACCTGAATCCATAGAGGAACGTGTGAATTTAAATAATGTTCTTCTCAATTAT 1113
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1114 CCATCAAGACCACTATCAAGATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAG 1173
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1174 ACAGTCGCTTGGTGGTCTCAATGGCAGTGGAGAGTACGGTAGTCCAGCTTCTGCAG 1233
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1234 AGGTTATATGATCCGATCATGGCTTTTATCATGGTGGATGAGAATCACATCAGAGCTTTA 1293
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1294 AATGTGGGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1353
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
DB 1354 ACCACCATCAGTAAACATATCAAGTATGGACGAGATGATGCTGACTGATGAGAGATGGAG 1413
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1414 ACATTTGGTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGGCGAAAAACAGAGGATCGCAATT 1473
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1474 ACATTTGGTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGGCGAAAAACAGAGGATCGCAATT 1533
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1534 GCTCGTGCCTTAGTTGCAAAACCCCAAGATTCTGATTTAGATGAGGCTAGCTCTGCCCTG 1593
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1594 GATTCAAGAAAGCAAGCTAGCTGTTCAAGCTGCACCTGGAGAAG 1635
```

## RESULT 3

US-09-873-409-16  
; Sequence 16, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:

; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1940  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-16

## Alignment Scores:

Pred. No.:	6,98e-270	Length:	1940
Score:	2568.50	Matches:	514
Percent Similarity:	95.01%	Conservative:	0
Best Local Similarity:	95.01%	Mismatches:	0
Query Match:	93.78%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-7 (1-541) x US-09-873-409-16 (1-1940)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
```

```
Db 34 ATGATCTCGGTATACCTGGCATCAGTGGTCAATGGAGCCCTGCCTCTCTTAATGCCACTG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 94 GTTTTAGAGAAATGAGTGATTAACCTTATTAGTGGATGTCTAGTCCAAATTAACATAC 153
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLeuLeuPheGly 60
Db 154 TCTTTCTTCAGTTGACCTGTATTATGTTGGATAGGTGTTGCTGCTTGATTTTGGT 213
Qy 61 TyrIleGlnIleSerLeuThrPheIleThrAlaAlaArgGlnThrIleArgIleArgLys 80
Db 214 TACATACAGATTTCTTGTGGATTAATACTGCAGCAGCAGACAGCAAGAGGATTCGAAA 273
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
Db 274 CAGTTTTTTTCATTGAGTTTGGCAGAGCATCGGCTGGTTGATAGCTGTGACATCGGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACCTTAACACTCGCATGACAGACATTCACAAATCAGTCATGTTGAGATAAGATT 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTTCGATTGGCCTGGCAGTTGGTGAAG 453
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCTGGAAACTCACCCTAGTAGCTCTATCCAGCGTCTCCTCTTATTAATGGCTTCAGCGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaThrLysAla 180
Db 514 GCATGTTCTAGGATGGTCACTCATTCACCCAGTAAGGAATTAAGTGCCTATTCNAAGCT 573
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGCTGTGGCAGAGAAGTCTTGTCATCAATCCGAACAGTCATAGCTTTAGGGCCAG 633
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db 634 GAGAAAGAACTTCAA----- 648
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 649 -----AGGTATACAGATCTCAAGAT 672
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 673 GCAAAAGGATTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTAC 732
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsn 280
Db 733 TTTCTTTATGAATGAACCTATGAGACTTGTCTTTTGGTATGAACCTCTTGTATCTTAAT 792
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 793 GGAGAACTGGATATACATCGGACTGTCTTGTCTTTTCTTACTGTATCCATAGC 852
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 853 AGTTATTGCAATGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCT 912
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 913 GCCTTTTCATATTTTCCAGGTATTGATAAGAAACCCAGTAGATAGATAACTTTTCCACAGCT 972
Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 973 GGATATAAACCCTGAAATCCATAGAGGAACCTGTGGAAATTTAAATGTTTCTTCAATAT 1032
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1033 CCATCAAGACCATCTATCAAGATTCTCAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAG 1092
```

```
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1093 ACAGTCGCCCTTGGTCGTCTCAATGCGACTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1213 AATGTCGCGCATTTATCGACACCATATTGGAGTGGTTAGTCAAGAGCTGTGTTTGTTCGGG 1272
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1273 ACCACCATCAGTAACATATCAAGTATGCGCAGATGATGTGACTGATGAGAGATGGAG 1332
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1333 AGAGCAGCAAGCAAGCAAAATCGTATGATTTTATCATGGAGTTTCTTAATAATTTAAT 1392
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1393 ACATTGTTAGGGGAAAGAGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAAT 1452
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCGATTTTAGATGAGGCTACGTCGCCCTG 1512
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 540
Db 1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAGGATACCCCCAGGTATTCA 1572
Qy 541 Phe 541
Db 1573 TTT 1575

RESULT 4
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.: 4,18e-265 Length: 3621
Score: 2528.50 Matches: 507
Percent Similarity: 94.94% Conservative: 0
Best Local Similarity: 94.94% Mismatches: 0
Query Match: 92.31% Indels: 27
DB: 10 Gaps: 1

US-09-873-409-7 (1-541) x US-09-873-409-14 (1-3621)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 34 ATGATCTCGGTATACCTGGCATCAGTGGTCAATGGAGCCCTGCCTCTCTTAATGCCACTG 93
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
```

Db 94 GTTTTAGGAGAAATAGTGATGAACCTTTATAGTGGAGTGTCTAGTCCAAACTAACACATAC 153  
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyValAlaAlaLeuIlePheGly 60  
Db 154 TCTTTCTTCAGGTGGACCTGTATATATGTTGAATAGTGTGCTGCTCTGATTTTGGT 213  
Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
Db 214 TACATACAGATTTCCCTGTGTGATTATTAACCTGCAGACGACGACAGCAAGAGGATTCGAAA 273  
Qy 81 GlnPhePheHiserValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
Db 274 CAGTTTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTGATAGCTGTGACATCGGT 333  
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 334 GAACTTTAACACTCGCATGCAGACAGACATTCAGAAATCAGTGATGGTATTCGAGATAAGATT 393  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 394 GCTCTGTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTGGTGAAG 453  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 454 GGCCTGAAAACCTCACCTAGTGACTCTATCCACGTCCTCTTATAATGGCTTCAGCGCA 513  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 514 GCATGTTCTAGGATGGTCACTCATCTGACCAAGTAAAGAAATTAAGTCGCTTATCCAAAGCT 573  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 574 GGGGCTGTGGCAGAAAGTCTTGTGTCATCAATCCGAACAGTCATAGCTTTAGGGCCAG 633  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 634 GAGAAAGAACTTCAA----- 648  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 649 -----AGGTATACAGAAATCTCAAGAT 672  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 673 GCAAAAGATTTTGGCATAAAGAGCTATAGCTTCAAAAGTGTCTCTGCTGTGTGTAC 732  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerIleIleLeuAsn 280  
Db 733 TTTCTTTATGAATGGAAACCTATGGACTTGTCTTTTGGTATGGAAACCTCTTGATTTCTTAAT 792  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 793 GGAAACCTGGATATACCATCGGAGCTGTTCTTGTGCTTTTCTTAGTGTAAATCCATAGC 852  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 853 AGTTATTGCTATGGACAGCAGCTCCCTCACTTTTGAACCTTCCGAATAGCCCGAGAGCT 912  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 913 GCCTTTTCATATTTCCAGGTTATTGTATGAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972  
Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 973 GGATATAAAACCTCAATCCATAGAGGAACCTGTGGAATTTTAAAGATGTTCTTTCAATTAT 1032  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1033 CCATCAAGACCATCTATCAAGATTTCAAGAGTCTGATCTCAGAAATTAAGTCTGGAGAG 1092  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1093 ACAGTGCCTTGGTCTCTCAATGGCAGTGGGAAGAGTAGTCGGTAGTCCAGCTTCTGCAG 1152

Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1153 AGGTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTA 1212  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1213 AATGTGCGGCATTTATCGAGACCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1272  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1273 ACCACCATCAGTAAACAATATCAAGTATGGACGAGATGATGTGACTGATGAGAGATGGAG 1332  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1333 AGAGCAGCAAGCAAGCAAAATCGTATGATTTTATCATCGAGTTTCTTAATAAATTAAT 1392  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1393 ACATTTGGTAGGGGAAAAAGGAGCTCAAAATGATGGAGGGCAGAAACAGAGGATCGCAATT 1452  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1453 GCTGTCCTTATGTTGAAACCCCAAGATCTGATTTTATAGATGAGGCTAGCTGCGCTG 1512  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1513 GATTCAAGAAAGCAAGTCAAGCTGTTCAAGCTGCACCTGGAGAAG 1554

## RESULT 5

US-09-873-409-12  
; Sequence 12, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 3177  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: No. US20020037522A1e  
; LOCATION: (198)..(198)  
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)  
US-09-873-409-12

Alignment Scores:  
Pred. No.: 2,978-193 Length: 3177  
Score: 1868.00 Matches: 370  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 68.20% Indels: 0  
DB: 10 Gaps: 0

US-09-873-409-7 (1-541) x US-09-873-409-12 (1-3177)

Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184  
Db 1 ATGGTCACTTCATTGACCAAGTAAAGAAATTAAGTGCCTATTCCAAAGCTGGCGCTGGCA 60  
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204  
Db 61 GAAGAAGTCTTGTTCATCAATCCGAACAGTCATAGCTTTAGGGCCCGGAGGAAGAACTT 120  
Qy 205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224

```
Db 121 CAAAGGCTTCCTTTTAAATATAACAAGATATGCTTGGTTTATTTTCCCGAGTGGCTA 180
Qy LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspPhe 244
Db 181 CTAAGTTGTCTGCTGTTNTTGTAAAGGTATACACAGAACTTCAAGAGTGCAGAGGATTTT 240
Qy GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
Db 241 GGCATAAAGAGGACTATAGCTTCBAAGGTCTCTTGGTCTGTGTACTTCTTTATGAAT 300
Qy GlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
Db 301 GGAACCTATGGACTTGTCTTTTGGTATGGAACCTCTCTTGATTTCTTAATGGAACCTGGA 360
Qy TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
Db 361 TATACCATCGGGACTGTCTTGTCTGTTTCTTTAGTGTAAATCCATAGCAGTTATGCAAT 420
Qy GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGTGCCTTTCATATT 480
Qy PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
Db 481 TTCAGGTTATTGATAAGAAACCCAGTATAGATAAATCTTTCCACAGCTGGATATAAACCT 540
Qy GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
Db 541 GAATCCATAGAAGAACTGTGGAAATTTAAATAATGTTCTTTCAATATATCCATCAAGACCA 600
Qy SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
Db 601 TCTATCAAGATTCTGAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCTTG 660
Qy ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
Db 661 GTCGGTCTCAATGGCAGTGGGAAGATACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720
Qy ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
Db 721 CCGGATGATGGCTTTATCATGGTGGATGGAGAAATGACATCAGAGCTTTAAATGTGCGGCAT 780
Qy TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
Db 781 TATCGACCATATTTGAGTGGTGTAGTCAAGAGCTGTGTTGTCGGACCAACCATCAGT 840
Qy AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
Db 841 AACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGG 900
Qy GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
Db 901 GAAGCAATCGTATGATATTTATCATGGAGTTTCCCTAATAAATTAATACATGTTGATGGG 960
Qy GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
Db 961 GAAAGAGAGCTCAATAGTGGAGGCGCAGAAACAGAGATCGCAATGCTGTCGCTTAA 1020
Qy ValArgAsnProTyrIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
Db 1021 GTTCGAAACCCCAAGATCTGATTTTATAGATGAGGCTACGCTGCTGCTGGATTCAGAAAGC 1080
Qy 525 LysSerAlaValGlnAlaLeuGluLys 534
Db 1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1110
```

## RESULT 6

US-09-866-866A-5  
; Sequence 5, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John

; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866, 866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584, 586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086, 988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 4189  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-866-866A-5

Alignment Scores:  
Pred. No.: 3,15e-160 Length: 4189  
Score: 1566.50 Matches: 307  
Percent Similarity: 72.63% Conservative: 99  
Best Local Similarity: 54.92% Mismatches: 94  
Query Match: 57.19% Indels: 59  
DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x US-09-866-866A-5 (1-4189)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTCCTGGAACTCTCGCTGCTATATATCATGGAACATTACTTCCCTCTTTGATGCTG 207
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAGCAGAACCCAGTAGTATTCGCCAAGCAATT 267
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAAAGTGGACCCACACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGGCCATATACGCCCTAC-----TATTACACCGGGATTGGTCTGCTGTGTG 372
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76
Db 373 CTATAGTTCCTTACATCCAGGTTTCATTTGGTCTGGCAGCTGGAAGACAGATACAC 432
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db 433 AAGATTAGGCAGAAAGTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492
Qy 97 CysAspIleGlyIleLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGCTCCAAAATTAATAGCGGA 552
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATTGGTGACAAAATGGGATGTTTTCAGTCCATCAACCAATTTTAGCCGGTTTATC 612
Qy 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGGATTTATAAGTGGTGGAAAGCTAACCTTGTTCATTTTGGCTGTCAGCCCTCTTATT 672
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATTGTCTCTCTTGTGGCAAAAGGTATTGACTTCACTTACTTAATAGGAACCTCCAG 732
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTTATGCAAAAGCTGGAGCAGTGTCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
```

Db 793 GCCTTTGGAGGCAACAGAGGAAGCTTGAA----- 822  
Qy 216 AlaTrpPheTyPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThr 235  
Db 823 -----AGGTACAAT 831  
Qy 236 GlnAsnLeuLysAspAlaLysAspPheGlyLeuArgThrIleAlaSerLysValSer 255  
Db 832 AAAAAATTAGAAAGCTAAAAATTGTGGCATAAAGAAAGCTATCACAGCCAGCATTTGG 891  
Qy 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275  
Db 892 ATAGGCATTGCCCTGTGTGCTATGCATCATATGCACATTCCTGGTATGGGACA 951  
Qy 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295  
Db 952 TCTTGTGCTCTCAATAGAA-----TATTCTATTGGAGAGTGTCTTACTGCTCTTC 1005  
Qy 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315  
Db 1006 TCTATTGTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCTTTGCA 1065  
Qy 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335  
Db 1066 AAGCGACAGGGCGACGCTTTGAAATCTTCAAGATAATTGATAACGAGCCACGATTGAC 1125  
Qy 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355  
Db 1126 AGCTTCTCAACAAGGGCTACAAACAGACAGTATATATGGAAACTTAGAGTTTAAAT 1185  
Qy 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375  
Db 1186 GTTCACTTCAACTACCATCGAGAGCGAAGTTTCAGATCTTGAAGGCTTCAATCTGAAG 1245  
Qy 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395  
Db 1246 GTGAAGAGCGGACAGACGGTGGCTTGGTTGGCAACAGTGGCTGTGGAAAAACACAAC 1305  
Qy 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsn 415  
Db 1306 GTCCACTGATGCAGAGGCTCTACGACCCCTCGAGGGCGTGTCTAGTATCGACGGACAA 1365  
Qy 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435  
Db 1366 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTTGTTGTGTGAGTGA 1425  
Qy 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455  
Db 1426 CCTGTGCTGTTGCCACCCAGCATCGCCGAGAACATTCGCTATGCCGAGAAAGATGTCACC 1485  
Qy 456 AspGluGluMetGluArgAlaAaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475  
Db 1486 ATGGATGAGATTCAGAAAGCTGTCAAGGAAGCCAAATGCCCTATCACTTCATCAAGA 1545  
Qy 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLys 495  
Db 1546 CCCACCATTTCACACCCCTGGTTGGTGAGAGAGGGCGCAGCTGAGTGGGGGACAGAA 1605  
Qy 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuLeuLeuAspGlu 515  
Db 1606 CAGAGAATGCCATTGCCCGGCGCTGTCCGCAATCCCAAGATCTCTTTGTTGGACGAG 1665  
Qy 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
Db 1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGTGCGAGCCGACCTGGATAAG 1722

RESULT 7

US-09-866-866A-1  
; Sequence 1, Application US/09866866A  
; Patent No. US20020102244A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian

; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 3860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-866-866A-1

Alignment Scores:

Pred. No.: 4.38e-159 Length: 3860  
Score: 1555.50 Matches: 306  
Percent Similarity: 73.29% Conservative: 100  
Best Local Similarity: 55.23% Mismatches: 99  
Query Match: 56.79% Indels: 49  
DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x US-09-866-866A-1 (1-3860)

Qy 1 MetIleLeuGlyIleIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 151 ATGGTGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 211 GTGTTGGAGAAATGACAGATATCTTTGCCAAATGCAGGAAATTTAGAAATCTGATGTCA 270  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe 43  
Db 271 AACATCCTAATAAGAAAGTATCATATCATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 330  
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 331 ATGACACAGATATGCCTATTATTATACATGGAAATTTGGTGGTGGTGGTGGTGGTGGT 390  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 391 ATTCAGGTTTATTTTGGTGGCTGGCAGCTGGAAGCAAAATACACAAAATTAGAAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGCTATAATGCAGCAGAGATAGCTGCTTTTGTATGTGCACGATGTTGGGGAG 510  
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 511 CTTTAACACCCGATCTACAGATATGTCTCTAAGATTAAATGAAGGTATTCGTGCACAAAT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGNATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 630  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GGTGGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCTCTTCTTGGAGTGTGAGTGTGT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerLysAla 180  
Db 691 GTCTGGGCAAGATATCTTCTTCAATTACTGATAAAGAACTCTTAGCTATGCAAAAGCT 750  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 811 AAGAAAGAACTTGAA----- 825  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 826 -----AGGTACACAAATAATTAGAGAA 849  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 850 GCTAAAGAAATGGGAATAAGAAAGATTACAGCCAAATTTCTATAGTGTCTGCTTTC 909  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsn 280  
Db 910 CTGCTGATCTATGCATTTATGCTCTGGCTTCTGGTATGGACCACTTGTCTCTCA 969  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 970 GGGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAATTGGG 1023  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1024 GCITTTAGTGTGGACAGGATCTCCAAAGCATTTGCAATGTCAGAGGAGCA 1083  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1084 GCITATGAATCTTCAAGATAATTTGATAAAGCAAAGTATTGACAGCTATTGGAAGAGT 1143  
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1144 GGGCACAACACAGATAATATTAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1204 CCATCTCGAAAAGAAATTAAGATCTTGAAGGCTTGAAGGCTGAAAGTGCAGAGTGGGCG 1263  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCAG 1323  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1324 AGGCTCTATGACCCACAGAGGGGATGCTAGTGTGATGACAGGATATTAGGACCATA 1383  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1384 AATGTAAGGTTTCTACGGGAAATCAITTTGGTGTGTGATGAGTACGGAACCTGATTGTTGCC 1443  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
Db 1444 ACCAGATAGCTGAAACATTCGCTATGGCCGTGAAATGTCAACATGGATGAGATTGAG 1503  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1504 AAAGCTCAAGGAAGCAATGCTATGACTTTATCATGAACTGCTCATAAAATTTGAC 1563  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIle 500  
Db 1564 ACCCTGTTGGAGAGAGGGGCCCACTTGAAGTGGTGGGCAAGCAGAGGATGCCATT 1623  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1624 GCAGTGGCCCTGGTTGCAACCCCAAGATCTCTCTGCTGGATGAGGACGCTGAGCCTTG 1683  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
Db 1684 GACACAGAAACGAAGCAGTGGTTTCAGTGGCTCTGGATAAG 1725

RESULT 8

US-10-044-671-1  
; Sequence 1, Application US/10044671  
; Patent No. US20020177147A1  
; GENERAL INFORMATION:

; APPLICANT: Washington State University Research Foundation  
; APPLICANT: Mealey, Katrina  
; APPLICANT: Bentjen, Steven  
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE  
; FILE REFERENCE: 4630-61733  
; CURRENT APPLICATION NUMBER: US/10/044,671  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/261,578  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: US 60/314,829  
; PRIOR FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4317  
; TYPE: DNA  
; ORGANISM: Canis familiaris  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (70)..(3912)  
; OTHER INFORMATION:  
US-10-044-671-1  
  
Alignment Scores:  
Pred. No.: 1,le-158 Length: 4317  
Score: 1552.50 Matches: 302  
Percent Similarity: 71.94% Conservative: 98  
Best Local Similarity: 54.32% Mismatches: 105  
Query Match: 56.68% Indels: 51  
DB: Gaps: 5  
  
US-09-873-409-7 (1-541) x US-10-044-671-1 (1-4317)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 217 ATGTTGGTGGGACAATGGCTGCCATCATCCATGGAGCTGCATCCCTCTCATGATGCTG 276  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 277 GTTTTGGAAACATGACAGATAGCTTTGCCAATGCAGGAATTTCAAGAAACAAACTTTT 336  
QY 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 337 CCAGTTATAATTAATGAAGTATTACGAACATATACCAACATTTTCATCAACCATCTGGAG 396  
QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
Db 397 GAGGAATGACCAAGTATGCTTATTATACAGTGGATCGGTGCTGGCGTGGTGGCT 456  
QY 60 GlyTyrIleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleArg 79  
Db 457 GCTTATCATCCAGTTTCATTCGTGCTGGCAGCAGGAGACAGATACATAAAATTAGA 516  
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99  
Db 517 AAACAATTTTTCATGCTATCATGCGACAGAGATGGCTGGTTTGCAGTGCATGACGTT 576  
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 577 GGGGAGCTTAACACCGGCTCACAGACGATGCTCAAAATCAATGAAGAAATGGCGAC 636  
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 637 AAGTTGGAATGTTCTTCATCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT 696  
QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 697 ACACCTGGTTGGAAGCTAACCCCTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTCA 756  
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 757 GCCGCCCATCGGCAAGATGACTATCTTCTTACTGATAAAGAACTCTTGGCGCTATGCA 816

QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
DB 817 AARGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 876  
QY 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218  
DB 877 GGACAAAGAAAGAACTTGA----- 897  
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeu 238  
DB 898 -----AGGTACAACAAATAATTGA 915  
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
DB 916 GAAGAAGCTAAAGAATTTGGGATAAAGAAAGTATCAGCGCAACATTTCTATTGGTGCC 975  
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278  
DB 976 GCTTTCTTATTGATCATGTCATATGCTCTGGCTTTCTGGTATGGACCTCTTGGTC 1035  
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
DB 1036 CTCTCCAGTGAA-----TATACTATTGGACAGTACTCACTGCTCTTCTTCTGTATTGA 1089  
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
DB 1090 ATTGGGGCTTTTAGTATTGGACAGGCATCCCAAGCAATTGAAGCAATTGCAACGCAAGA 1149  
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
DB 1150 GGAGCAGCTTATGAATCTTCAAGATAAATTTGCAATAAACCAGCAAGTACACACTATTGC 1209  
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358  
DB 1210 AAGAGTGGACATAAACAGATATATTAGGGAATTTGGAATTCAAATGTTCACTTTC 1269  
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
DB 1270 AGTTACCTCTTCGAAAAGAAAGTAAAGATCTTAAGGGTCTCAACCTGGAAGGTTACAGAT 1329  
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
DB 1330 GGCCACACAGTGGCGTGGTGGGAACAGTGGCTGGGGAAGACACGACCGGTGCAGCTG 1389  
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
DB 1390 ATGCAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGACAGGACATTAGG 1449  
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
DB 1450 ACCATAAATGTAAGGCATCTTCGGGAAATTTACTGTTGTGTGTGAGTCAGGAGCCTGTGTTG 1509  
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458  
DB 1510 TTTGCCACCACGATAGCTGAAAACATTTGCTATGGCCGCGCAAAATGTCTACCATGATGAG 1569  
QY 459 MetGluArgAlaAlaArgGluAlaAenAlaTyrAspPheIleMetGluPheProAsnLys 478  
DB 1570 ATTGAGAAAGCTGTTTAGGAAGGCAATGCTCTATGATTTTATCATGAATCTACCTAATAAA 1629  
QY 479 PheAsnThrLeuValGlyGlyLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
DB 1630 TTTGACACTCTGTTGGAGAGAGAGGGCCGCGCTGAGTGGTGGACAGAAACAGAGAAATC 1689  
QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
DB 1690 GCCATTGCTCGGGCCCTGGTTGCGCAACCCCAAGATTTCTTGTGATGATGAGCAACGTCGA 1749  
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
DB 1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGGCCCTGGATAG 1797

RESULT 9

US-09-917-800A-1424  
; Sequence 1424, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1424  
; LENGTH: 4254  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855

US-09-917-800A-1424

Alignment Scores:  
Pred. No.: 3,34e-158 Length: 4254  
Score: 1548.00 Matches: 307  
Percent Similarity: 72.25% Conservative: 94  
Best Local Similarity: 55.32% Mismatches: 104  
Query Match: 56.52% Indels: 50  
DB: 10 Gaps: 5

US-09-873-409-7 (1-541) x US-09-917-800A-1424 (1-4254)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 247 ATGGCTCTGGGAAGTCTCGCTGCTATCATCCACGGAACCTGCTTCCCTCTCGATGCTG 306  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
DB 307 GTGTTCGGATACATGACAGATAGTTTACCCCAAGCAGAGACCCGCAATCTGACCGAGCG 366  
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45  
DB 367 ATTTACTAATCAAAAGTGAATCAACAGTACACATACATACCTCAGCGCACAGAGTCTGGAGGAG 426  
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
DB 427 GACATGCCCATGTATGCCCTACTATTATACGGGCAATTTGGTCCCGGTGTGCTCATCTGTGCC 486  
QY 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
DB 487 TACATCCAGGTTTCACTTTGGTCCCTGGCAGCTGGGAGACAAATACACAGATTATGGCAG 546  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100



[illegible]

```
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATGCCAGGAGATAGGCTGTTTCATGTGCACCATGTTGGGGAG 510
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAAACACCCGCTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTTGGTGCAAAATT 570
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGCTTTTCAGTCATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 630
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGTGAAGCTAACCTTGTGATTTTGGCAGCATCAGTCTGTTCTTGGACTGTGCAGCTGCT 690
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCTGGCAAGATATCTATCTTCTATTACTGATTAAGAACTCTTAGCGTATGCAAAAGCT 750
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTTAGAACTGTGATTGCAATTGGAGGACAA 810
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
Db 811 AGAAAGAACTTGAA----- 825
QY 221 ProGlnTyrLeuLeuSerCysValLeu---PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 826 -----AGGTACAAACAATAATTAGAAAGAA 849
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 850 GCTAAAGAAATTCGGATTAAGAAAGCTATTACAGCAATTTCTATAGTGCTGCTTTC 909
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuLeuAsn 280
Db 910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTTGGTCTCTCA 969
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 970 GGGGAA-----TATTCTATTGGCAAGTACTCACTGTATTCTTTCTGPTAATTTGGG 1023
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1024 GCATTATGAAATCTTCAAGATAATTGATAATAAGCCAAAGTATTGACAGCTATTGCAAGAGT 1143
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1084 GCATTATGAAATCTTCAAGATAATTGATAATAAGCCAAAGTATTGACAGCTATTGCAAGAGT 1143
QY 341 GlyTyrLysProGluSerIleGlyGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1144 GGGCACAACACAGATAATATTAAGGGAAATTTGGAATTTCAAGAAATGTTCACTTTCAAGTAC 1203
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1204 CCATCTCGAAAGAAAGATTAGATCTTGAAGGGCTGAACTGAAGGTGCAGATGGGCAG 1263
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1264 ACGGTGGCCCTGTTGGAAACAGTGGCTGTGGGAAGCAGACACAGTCCAGCTGATGCAG 1323
QY 401 ArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1324 AGGCTCTATGACCCCAAGAGGGGATGGTCACTGCTGATGCAGAGGATATTAGGACCATTA 1383
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1384 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGTGAGTCAGAACCTGATTGTTGTC 1443
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
```

```
Db 1444 ACCCAGATAGCTGAACAAACATTCGCTATGGCCGTGAATAATGTCCACCATGATGAGATTGAG 1503
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1504 AAAGCTGTCAAGAGCCCAATGCCCTATGACTTATCATGAATAACTGCCTCATAAATTGAC 1563
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1564 ACCCTGCTGGAGAGAGGGGCCAGTTGAGTGTGGTGGCAGAGCAGAGGATCGCCATT 1623
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1624 GCACGTGCCCTCGTTCGCAACCCCAAGATCCTCTGCTGGATGAGGCCACGTCAGCCTTG 1683
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1684 GACACAGAAAGCGACGATGGTTCAGGTGGCTCTGGATAAG 1725

RESULT 11
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:
Pred. No.: 1,466-157 Length: 8630
Score: 1546.50 Matches: 305
Percent Similarity: 73.10% Conservative: 100
```

Best Local Similarity: 55.05% Mismatches: 100  
Query Match: 56.46% Indels: 49  
DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x US-09-306-417-1 (1-8630)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGTGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AnLeuIleSer 31
DB 1430 GTGTTGGAGAAATACAGATATCTTTGCAAAATGCAGAAATTTAGAGATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyr-SerPhePhe----- 43
DB 1490 AACATCTAATAGNAGTATCAATGATACAGGGTTCCTCATGATCTGGAGGAAGAC 1549
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACACAGATATGCTATTATACAGTGAATGGTCTGGGGTGGCTGGCTGTCTTAC 1609
QY 62 IleGlnIleSerLeuThrIleLeuAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 1610 ATTCAGGTTTCATTTGGTGGCAGTGAAGACAAATACACAAATATAGAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheLeuSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTAATGCGACAGGATAGGCTGGTTGATGTCACGAATGGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 1730 CTTAACACCCGACTTACAGATGATCTCTAAGATTAATGAAGTTATGGTGACAAAT 1789
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 1790 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1850 GGTTGGAGCTTAACCTTGATTTTGGCATCATGCTGCTGTTCTGGACATGTCAGCTGCT 1909
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1910 GTCTGGCGCAAGATACTATCTTCTTACTATGAAGAAGCTCTTAGCGTATGCCAAGCT 1969
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1970 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAAGCTGTGATGCAATTTGGAGGACAA 2029
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 2030 AAGAAAGAACTTGAA----- 2044
QY 221 ProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 2045 -----AGGTACAACAAAATTTAGAGAA 2068
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 2069 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGTGTGCTCTTC 2128
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 2129 CTGCTGATCTATGATCTTATGCTCTGGCCCTTCTGGTATGGGACCACTTGGTCTCTCA 2188
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 2189 GGGGAA-----TATTCTATTGGCAGACTACTGATTTCTTTCTGATTAATTGGG 2242
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 2243 GCTTTTAGTGTGGACAGGCACTCCAAAGCATTTGCAAGCATTTGCAAAATGCAAGAGGACGA 2302
```

```
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAAGTATTGACAGCTATTCCGAAGAGT 2362
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 2363 GGGCAAAACCAAGATATAATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 2423 CCATCTCGAAAAGAGTTAAGATCTTGAAGGGCTGAACTGAAGTGCAGAGTGGGCGAG 2482
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 2483 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACACAGTCCAGCTGTATGCAG 2542
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 2543 AGGCTCTATGACCCACAGAGGGGATGGTCAAGTGTGATGGACAGGATATTAGGACCAT 2602
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 2603 AATGTAAGGTTTCTACGGGAAATCATTTGGTGTGGTCAAGGAACTGTATTGTTGTC 2662
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
DB 2663 ACCACGATAGCTGAAACCATTCGCTATGCGCGTGAAATGTCCACCATGGATGAGATTGAG 2722
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 2723 AAGCTGTCAAGGAAGCAATGCTATGATCTTATCATGAAACTGCTCATAAATTTGAC 2782
QY 481 ThrThrValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 2783 ACCCTGGTTGGAGAGAGGGGGCCAGTGTGAGTGGGGGAGAGGAGGATCGCCATT 2842
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 2843 GCAGTGGCTCGTGGTTCGCAACCCCAAGATCCCTCTGATGAGGCCACAGTTCAGCTTG 2902
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
DB 2903 GACACAGAAAGCGAAGCAGTGGTTTCAAGTGGCTCTGGATAAG 2944
RESULT 12
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta91msAl
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
```

OTHER INFORMATION: plasmid backbone (pUC)

FEATURE:

NAME/KEY: misc feature

LOCATION: (161)..(677)

OTHER INFORMATION: 5'-LTR

FEATURE:

NAME/KEY: 5'UTR

LOCATION: (532)..(1219)

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: (1220)..(5062)

OTHER INFORMATION: msal mdrl cdna

FEATURE:

NAME/KEY: misc feature

LOCATION: (5215)..(5774)

OTHER INFORMATION: 3'-LTR

FEATURE:

NAME/KEY: misc feature

LOCATION: (5775)..(8630)

OTHER INFORMATION: plasmid backbone (pUC)

US-09-306-417-2

Alignment Scores:

Pred. No.: 1,46e-157 Length: 8630  
Score: 1546.50 Matches: 305  
Percent Similarity: 73.10% Conservative: 100  
Best Local Similarity: 55.05% Mismatches: 100  
Query Match: 56.46% Indels: 49  
DB: 6 Gaps: 6

US-09-873-409-7 (1-541) x US-09-306-417-2 (1-8630)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 1370 ATGGTGGGGAACCTTGGTGCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 1430 GTGTTTGGGAATGCAGATATCTTTGCAAAATGCAGAAATTTAGAAATCTGATGTCA 1489  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 1490 AACATCACTAATAGNAGTGATATCAATGATACAGGGTCTTCTCATGATCGAGAGAC 1549  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 1550 ATGACAGGATGCTTATTAATACAGTGAATTTGGTCTGGGTGCTGGTGTCTGCTTAC 1609  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 1610 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAGACAAATACACAAATTTAGAAACAG 1669  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 1670 TTTTTCATGCTATAATGCACAGGATAGGCTGGTTTTCATGTCACGATGTTGGGGAG 1729  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 1730 CTTAACACCCGATACAGATGATGCTCTTAAGATTAATGAAGTATTGGTGACAAATTT 1789  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 1790 GGAATGTTCTTCAGTCAATGGCAACATTTTTCATCGGTTTATAGTAGGATTTACACGT 1849  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 1850 GGTGGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCTCTTCTTGGCATGTCAGCTGCT 1909  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180  
DB 1910 GTCTGGGCAAGATACATCTCTTACTTACTAATAAGAACTCTTAGCGTATGCAAAAGCT 1969  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

DB 1970 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATTCATTGGAGACAA 2029  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 2030 AAGAAGAACTTGAA----- 2044  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 2045 -----AGGTACAAACAAAAATTTAGAGAA 2068  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 2069 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGGTGCTGCTTC 2128  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
DB 2129 CTGCTCATCTATGCATCTTATGCTCTGGCTTCTGGTATGGACACCTTGGTCTCTCA 2188  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 2189 GGGAA-----TATTCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAATTGGG 2242  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 2243 GCTTTTAGTGTGGACAGGCATCTCCAAAGCATTTGAAGCATTTGCCAAATGCAAGAGGAC 2302  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 2303 GCTTATGAATCTTCAAGATATTTGATATAGCCCAAGTATTGACAGCTATTCCGAAGGT 2362  
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 2363 GGGCACAAACCAACATAATATTAAAGGAAATTTGGAATTTGAGAAATGTTCACTTCAGTTAC 2422  
QY 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 2423 CCATCTCGAAAAAGATTGAAGATCTTGAAGGGGCTGAACTGAAAGTGCAGAGTGGGCAG 2482  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 2483 ACAGTGGCTCTGGTGGAAACAGTGGCTGGGAGAGACACACAGTCCAGCTGATGTCAG 2542  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspLysAsnAspIleArgAlaLeu 420  
DB 2543 AGGCTCTATGACCCACACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACCAT 2602  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 2603 AATGTAAAGTCTTACGGGAAATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTGCC 2662  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 2663 ACCAGTACGTGAAACCAATTCGTATGGCCGTTGAATATGTACCATGGATGATGATGAG 2722  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 2723 AAAGCTGTCAAGGAAGCAATGCTTATGATCTTATCATGAAACTGCTCATAAATTTGAC 2782  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 2783 ACCCTGTTGGAGAGAGAGGGCCCAATTTGAGTGGTGGGAGAGAGAGATCGCCATT 2842  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCTCTGATGAGGCCACGTCAGCCTTG 2902  
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
DB 2903 GACACAGAAAGCGACGATGGTTTCAGTTGGCTCTCGATAAG 2944  
RESULT 13  
US-09-769-097-3

```

; Sequence 3, Application US/097609097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/207770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.: 1,24e-157 Length: 4425
Score: 1543.00 Matches: 300
Percent Similarity: 73.04% Conservative: 101
Best Local Similarity: 54.64% Mismatches: 102
Query Match: 56.33% Indels: 46
DB: 10 Gaps: 5

US-09-873-409-7 (1-541) x US-09-769-097-3 (1-4425)
Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 499 ATGCTGTGGGAACCTCTGGCGCCATTATCCATGGAATGGCGCTCCCACTTATGATGCTG 558
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 559 GTCTTTGGAGACATGACAGATAGCTTTGCAAT-----GTAGGAAACAAACCGTAGTAGT 612
Qy 41 SerPhePheArgLeuThr----- 46
Db 613 AGTTTCTCAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACACCGTAGCGCC 672
Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 673 TACTATTACAGGCAATTTGGTCCCGTGTCTCATCGTTGCCATCATCCAGGTTTCCACT 732
Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 733 TGGTGCCTGGCAGCTGGGAGACAAATACACAGATAGGACGAAAGTTTTTCCATGCCATC 792
Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGATAGCTTTGGGAGCTCAACACCCGGCTC 852
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db 853 ACAGATGACGCTCCAAATTAATGAAGAAATTTGTGACAAATTTGGAATGTTCTTTCAG 912
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrTrpLysLeuThr 145
Db 913 GCAATGGCAACATTTTTTGGTGGTTTTATATAGGATTTACTCGCGCTGGGAAGCTAACT 972
Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165
Db 973 CTGTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCTGAGCTGTTTGGCCAAAGATA 1032

```

Db 2026 GCGGTGTTTCAGCGCGCTCTGGATAG 2052  
::: ||||||| ||||||| ||||||| |||||||

## RESULT 14

US-09-769-097-1

; Sequence 1, Application US/09769097

; Patent No., US20020055128A1

; GENERAL INFORMATION:

; APPLICANT: Kimberly Anne Brun

; APPLICANT: Richard James Chenery

; APPLICANT: Harna Ellens

; APPLICANT: John Anthony Feild

; APPLICANT: Lin Yue

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES

; TITLE OF INVENTION: ENCODING RAT MOR1A AND SCREENING METHODS THEREOF

; FILE REFERENCE: GP-50009-C2

; CURRENT APPLICATION NUMBER: US/09/769, 097

; CURRENT FILING DATE: 2001-01-25

; PRIOR APPLICATION NUMBER: 09/208, 809

; PRIOR FILING DATE: 1998-12-09

; PRIOR APPLICATION NUMBER: 09/156, 800

; PRIOR FILING DATE: 1999-09-17

; PRIOR APPLICATION NUMBER: US99/20770

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 4369

; TYPE: DNA

; ORGANISM: RATTUS RATTUS

US-09-769-097-1

## Alignment Scores:

Pred. No.:	2, 01e-157	Length:	4369
Score:	1541.00	Matches:	300
Percent Similarity:	73.04%	Conservative:	101
Best Local Similarity:	54.64%	Mismatches:	102
Query Match:	56.26%	Indels:	46
DB:	10	Gaps:	5

US-09-873-409-7 (1-541) x US-09-769-097-1 (1-4369)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	499	ATGCTGCTGGAACTCTGGCGGCATATTCATGGAATTCGCTCCACATTCATGATGCTG	558
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	559	GTCTTTGGAGACATGACAGATAGCTTTTGCAAT-----GTAGGAAACACACCGTAGTATG	612
Qy	41	SerPhePheargLeuThr-----	46
Db	613	AGTTTCTACAAATGCTACAGACATATATGCAAGCTGGAGGACGAATATGGCCAGTACGCC	672
Qy	47	LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu	66
Db	673	TACTATTACGGGCAATGTTGGCGGTGCTCATCGTTCCTACATCCAGGTTTCATT	732
Qy	67	TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal	86
Db	733	TGTTGCTGCGAGCTGGGAGACAAATACACAGATTAGGCAGAAATTTTCCATGCCATC	792
Qy	87	LeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMet	106
Db	793	ATGAATCAGGAGATAGGCTGTTTGACGTGCATGACGTTCGGGAGCTCAACACCCGGCTC	852
Qy	107	Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln	125
Db	853	ACAGATGACGCTCCCAAAATTAATGAAGGAATTTGGTGAACAAATTTGGAATGTTCTTTTCAG	912
Qy	126	AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrPheLeuThr	145
Db	913	GCAATGGCAACATTTTTTTTGGTGGTTTTTAATAGGATTTACTCGCGGCTGGAAGCTAACT	972

Qy	146	LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAalaCysSerArgMet	165
Db	973	CTTGTAATTTGGCCATCAGCCCTGTTCTTGACCTGTAGCTGGTATTTGGGCAAGATA	1032
Qy	166	ValIleSerLeuThrSerLysGlyLeuSerAlaValSerLysAlaGlyAlaValAlaGlu	185
Db	1033	TTGTCTTCAATTTACTGATAAGGAATCCAGGCTTATGCAAAAGCTGGAGCAATGCTGAA	1092
Qy	186	GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln	205
Db	1093	GAAGTCTTAGCAGCCATCAGAATCTGATTCCTTTGGAGGACAAAAGAAAGCACTTGA	1152
Qy	206	ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu	225
Db	1152	-----	1152
Qy	226	SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly	245
Db	1153	-----AGGTACAATAAATTTGGAAGAGCTAAAGGCTTTGGG	1191
Qy	246	IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly	265
Db	1192	ATAAGAAAGCTATCACGCCCAACATTTCCATGGGTGAGCTTTCTGCTTATCTATGCA	1251
Qy	266	ThrTyrGlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr	285
Db	1252	TCATATGCTCTGCAATCTGTTATGGACTTCTCTGGTCTCATCTCAAAAGAA-----TAC	1305
Qy	286	ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly	305
Db	1306	ACTATTGGACAAGTCTCACTGCTTTTTTCTGTATTATTTGGAGCATTCAGTGTGGG	1365
Qy	306	AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe	325
Db	1366	CAGGCATCTCCAAATATTGAAGCTTCGCAATGCTAGAGGAGCAGCTTATGAAGTCTTC	1425
Qy	326	GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu	345
Db	1426	AGTATAATTTGATAATAAAGCCAGTATAGACAGCTTCTCAAGAGTGGGCACAAACCGAC	1485
Qy	346	SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer	365
Db	1486	AACATACAGAAATTTGGAATTTCAAAAATATTTCATCTCAGTTACCCGCTCGAAAGAC	1545
Qy	366	IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal	385
Db	1546	GTTCAGATCTTGAGGCGCTCAACCTGAAGGTGAAGAGCGGCGAGACGGTAGCCCTGTT	1605
Qy	386	GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspPro	405
Db	1606	GGCAACAGTGGCTGTGGGAAAGACACAACTGTCAGCTGTCGAGAGGCTCTACGACCCC	1665
Qy	406	AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr	425
Db	1666	ATAGAGGGCGAGTCACTAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG	1725
Qy	426	ArgAspHisIleGlyValSerGlnGluProValLeuPheGlyThrThrIleSerAsn	445
Db	1726	CGGAAATCATTTGGGTGGTGGTTCAGGAAACCGTCTGTTGCCACCAATTCGCGAA	1785
Qy	446	AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu	465
Db	1786	AACATTCGCTATGGCCGAGAAACAGTCACCATGTAGTATAGAGAAGCTGTCAAGAA	1845
Qy	466	AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu	485
Db	1846	GCCAAATGCTTATGATTTTCATGAAACTGCCCCCAAAATTTGCACACCTGTTGGTGAG	1905
Qy	486	LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal	505
Db	1906	AGAGGGCGCAGCTGATGGGGACAGAAACAGAGATTCGCTCCCGGCGCTGGT	1965

```
Qy 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db 1966 CGCAACCCCAAGATCCCTTTGGATGAGGCGCCACGTCAGCTTGGACACAGAAAGCGAA 2025

Qy 526 SerAlaValGlnAlaLeuGluLys 534
Db 2026 GCCGTGGTTCAGGCGGCTCTGGATAAG 2052

RESULT 15
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:
Pred. No.: 3,65e-157 Length: 4643
Score: 1539.00 Matches: 305
Percent Similarity: 73.10% Conservative: 100
Best Local Similarity: 55.05% Mismatches: 99
Query Match: 56.19% Indels: 50
DB: 9 Gaps: 7

US-09-873-409-7 (1-541) x US-10-072-621-2 (1-4643)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACCTCCCTCATGATGCTG 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAAATGACAGATATCTTCCAAATGCAGAAATTTAGAGATCTGATGCA 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAGTATCAATCATACAGGGTTCTTTCATGAATCTGGAGGAAGAC 754

Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCGATGATGCTATATTATACAGTGGAAATGGTCTGGGGTCTGGTCTGCTTAC 814

Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTACAGTTTCATTTGGTCTGGCAGCTGGAGACAAATACACAAATTTAGAAACAG 874

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATCGCAGCAGGATAGCTGGTTTGTATGTGCACGATGTTGGGAG 934

Qy 102 LeuAsnThrArgMetThr-----AspIleAspIleSerAspGlyIleGlyAspIle 120
Db 935 CTTRACACCCGACTTACAGATGATGCTTCCCAAGATTAATGAAGGAATTTGGTGACAAAT 994

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGGCACATTTTTCATCTGGTGGTTATAGTAGGATTTACAGT 1054

Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 160
Db 2045 GCACGTGCTGGTTCGCAACCCCAAGATCTCTCTGATGAGGCCACGCTCAGCCTTG 2104
```

```
Db 1055 GGTGGAAGCTAACCCCTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGCTGCTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACCTTCTTACTGTATAAAGAACTCTTAGCGTATGCAAGACT 1174
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGTGATTGCATTTGGAGGACAA 1234
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
Db 1235 AAGAAAAGAACTTCAA----- 1249
Qy 221 ProGlnTyrLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 1250 -----AGGTACAACAAAATTTAGAAAGAA 1273
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 1274 GCTAAAAGAAATTTGGATAAAGAAAGCTATTACAGCCAATATTCTATAGGTGCTGCTTTC 1333
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerIleLeuLeuAsn 280
Db 1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGACCACTTGGTCTCTCA 1393
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1394 GGGAA-----TATCTATTTGACAAGTACTCAGTGA---TTTTCTGTATTAATTTGGG 1444
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1445 GCTTTTAGTCTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGACGA 1504
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db 1505 GCTTATGAAATCTTCAAGATAATTTGAATGAAGCAAGTATTCACACTATTCGAAGAGCT 1564
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db 1565 GGGCACAACACAGATATATTAAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1624
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db 1625 CCATCTCGAAAAGAAAGTAAAGATCTTCAAGGGTCTGAACCTGGAAGTGCAGAGTGGGCAG 1684
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db 1685 ACGTGGCCCTGGTTGGAAAACAGTGGCTGGGGAAGACACACAGTCCAGCTCATGTCAG 1744
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1745 AGGCTCTATGACCCACAGAGGGGATGCTCAGTGTGATGGACAGGATATTAGGACCATA 1804
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1805 AATGTAAGTCTTACGGGAAATCATTTGGTGTGTGAGTCAAGAACCTGTATTGTTGCC 1864
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460
Db 1865 ACCACGATATGAAAACATTCGCTATGGCGGTGAAATGTCACCATGGATGAGATGAG 1924
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1925 AAAGCTCTCAAGGAAGCAATGCTATGACTTTATCATGAAGTCCCTCATAAATTTGAC 1984
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db 1985 ACCCTGTTGGAGAGAGAGGGGCCAGTTGATGTTGGTGGCAGAGAGGATGCCATT 2044
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2045 GCACGTGCTGGTTCGCAACCCCAAGATCTCTCTGATGAGGCCACGCTCAGCCTTG 2104
```



Search completed: March 31, 2003, 16:14:26  
Job time : 261.92 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 7313.73 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739

Sequence: 1 MILGILASLVNGACPLMLP.....SEKSAVQALEKDPTRYSF 541

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003.115418.19228/app\_query.fasta.1.7544  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -DOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 54887 @runat\_27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb.ba.\*  
2: gb.hcg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.ste.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.ste.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.bv.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2737	99.9	2021	6	AX339033 Sequence
2	2697	98.5	3702	6	AX339031 Sequence
3	2568.5	93.8	1940	6	AX339034 Sequence
4	2528.5	92.3	3621	6	AX339032 Sequence
5	2437	89.0	3699	6	AX478104 Sequence
6	1868	68.2	3177	6	AX339030 Sequence
7	1566.5	57.2	4189	6	AX322791 Sequence
8	1566.5	57.2	4298	10	MUSMDR
9	1561.5	57.0	4279	6	AX105082 Sequence
10	1557.5	56.9	4279	6	AX105078 Sequence
11	1557.5	56.9	4279	6	AX105080 Sequence
12	1557	56.8	4195	6	AX108656 Sequence
13	1556.5	56.8	4045	12	AF269224
14	1555.5	56.8	3860	6	AX322787 Sequence
15	1555.5	56.8	4378	6	E02326
16	1555.5	56.8	4646	6	AX391099 Sequence
17	1555	56.8	4280	10	CRUPGPII
18	1552.5	56.7	3888	10	AY082609
19	1552.5	56.7	4233	6	AR123273
20	1552.5	56.7	4279	6	AX105057 Sequence
21	1552.5	56.7	4317	4	AF045016
22	1552.5	56.7	4317	6	AX105059 Sequence
23	1552.5	56.7	4669	6	AR091275 Sequence
24	1552.5	56.7	4669	6	AR203322 Sequence
25	1548.5	56.5	4186	6	AX108654 Sequence
26	1548	56.5	4254	6	AX401748 Sequence
27	1548	56.5	4254	10	RATMDRM
28	1548	56.5	4927	10	AF257746
29	1547	56.5	4323	10	AF286167
30	1546.5	56.5	3860	6	AX322789 Sequence
31	1546.5	56.5	4646	6	AX336420 Sequence
32	1546.5	56.5	4646	6	AX336708 Sequence
33	1546.5	56.5	4646	6	I49610
34	1546.5	56.5	4646	9	HUMMDR1
35	1546.5	56.5	4669	6	I08557 Sequence 3
36	1546.5	56.5	6505	6	AR028671 Sequence
37	1546.5	56.5	8630	6	AX012320 Sequence
38	1546.5	56.5	8630	6	AX012321 Sequence
39	1546.5	56.5	9318	6	AR028672 Sequence
40	1545	56.4	4264	6	AR051647 Sequence
41	1545	56.4	4264	6	AR051650 Sequence
42	1542	56.3	4356	10	MUSMDRAA
43	1541.5	56.3	4296	10	CRUPGPI
44	1540.5	56.2	4304	10	CRUPGPI165
45	1540	56.2	3489	4	AB029153 Felis cat

ALIGNMENTS

RESULT 1

AX339033  
LOCUS AX339033 2021 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 15 from Patent WO0194400.  
ACCESSION AX339033  
VERSION AX339033.1 GI:18129125  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 15 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
source Location/Qualifiers  
1. 2021  
BASE COUNT 547 a 406 c 462 g 605 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,61e-214 Length: 2021  
Score: 2737.00 Matches: 541  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.93% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-7 (1-541) x AX339033 (1-2021)  
QY 1 MetLeuGlyLeuLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 34 ATGATCTGGGTATACCTGGCATCTGGTCAATGGAGCCCTGCCCTTCCTTAATGCCCATG 93  
QY 21 ValLeuGlyLeuMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 94 GTTTTAGAGAAATGAGTATGATTAATAGTGGATGCTAGTCCAAACTAACACATAC 153  
QY 41 SerPhePheArgLeuThrLeuTyrValGlyLeuValAlaLeuLeuPheGly 60  
DB 154 TCCTTCTTCAGGTGACCCGTATATGTTGGATAGGTGTTGCTGCTTGAATTTTGGT 213  
QY 61 TyrileGlnileSerLeuTyrPheileThrAlaAlaArgGlnThrIlysArgileArgLys 80  
DB 214 TACATACAGATTTCCCTTGTGGATTATTAACATGACGACGACGACGAGGATTCGAAAA 273  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspileGlyTrpPheAspSerCysAspileGly 100  
DB 274 CAGTTTTTTTCATTCAGTTTTGGCACGACACATCGGCTGGTTGTAGTACTGTGACATCGGT 333  
QY 101 GluLeuAsnThrArgMetThrAspileAspLysileSerAspGlyileGlyAspLysile 120  
DB 334 GAACCTTAACACTCGCATGACAGACATTCACAAAATCAGTCATGGTATTCGAGATAAGATT 393  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerileGlyLeuAlaValGlyLeuValLys 140  
DB 394 GCTCTGTTGTTTCAAAACATGCTACTTTTCGATTGGCCCTGGCAGTGTGGTGGTGAAG 453  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuileMetAlaSerAlaAla 160  
DB 454 GGCTGGAAACTCACCTTAGTGACTCTATCCACGTCCTCTCTTATATATGGCTTCAGCGGCA 513  
QY 161 AlaCysSerArgMetValileSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 514 GCATGTTCTAGGATGGTCACTCTCATTCAGCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573  
QY 181 GlyAlaValAlaGluValLeuSerSerileArgThrValileAlaPheArgAlaGln 200  
DB 574 GGGGCTGTGGCAGAAAGTCTTGTCATCAATCCGAAACAGTCATAGCCTTTAGGGCCAC 633

QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnileThrArgTyrAlaTrpPheTyrPhe 220  
DB 634 GAGAAAGAACTTCAAGGCTCTTCTCTTTTAAATATAACAAGATATGCTTGGTTTATTTT 693  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 694 CCCAGTGGCTACTAGTTGTGTCTGTTTGTAAAGGTATACACAGAATCTCAAGAT 753  
QY 241 AlaLysAspPheGlyLeuLysArgThrileAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 754 GCMAAGATTTTGGCATAAAGAGCTATAGCTTCAAAAGTGTCTCTTGGTGTCTGTGCTG 813  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuLeuLeuAsn 280  
DB 814 TTCTTTATGAATGAACCTATGGAACCTTCTTCTTGTATGGAACCTCTTCTTGAATCTTAAT 873  
QY 281 GlyGluProGlyTyrThrileGlyThrValLeuAlaValPhePheSerValileHisSer 300  
DB 874 GGAGAACCTGGATATACCATCGGACTGTCTTCTTGTCTTCTTCTTGTATTAATCCATAGC 933  
QY 301 SerTyrCysileGlyAlaAlaValProHisPheGluThrPheAlaileAlaArgGlyAla 320  
DB 934 AGTTATGCAATGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGAGCT 993  
QY 321 AlaPheHisilePheGlnValileAspLysLysProSerileAspAsnPheSerThrAla 340  
DB 994 GCCTTTTCATATTTCCAGGTATTGATAAGAAACCCAGCATATAGATAACTTTCCACAGCT 1053  
QY 341 GlyTyrLysProGluSerileGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 1054 GGATATAAACCTGAATCCATAGAGGAACCTGTGGAAATTTAAATAATGTTCTTTCAATTAT 1113  
QY 361 ProSerArgProSerileLysileLysileLysileLysileLysileLysileLysile 380  
DB 1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAAATCTCAGAAATTAAGCTCGAGAG 1173  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1174 ACAGTCGCTTGGTGGTCTCAATGSCAGTGGGAGAGTACGCTAGTCTCCAGCTTCTCGAG 1233  
QY 401 ArgLeuTyrAspProAspAspGlyPheileMetValAspGluAsnAspileArgAlaLeu 420  
DB 1234 AGGTTATATGATCCGATGATGGCTTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293  
QY 421 AsnValArgHisTyrArgAspHisileGlyValValSerGlnGluProValLeuPheGly 440  
DB 1294 AATGTCGGCATTAUCAGACCAATATGAGTGGTGTAGTCAAGAGCTGTTTGTTCGGG 1353  
QY 441 ThrThrIleSerAsnAsnileLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
DB 1354 ACCACCATCAGTAACAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAG 1413  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheileMetGluPheProAsnLysPheAsn 480  
DB 1414 AGAGCAGCAAGGAGCAATGCGTATGATTTTATCATGGAGTTTCCCTAATAAATTAAT 1473  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgileAlaile 500  
DB 1474 ACATTTGGTGGGAAAAGAGCTCAAAATGAGTGGAGGCGAGAAACAGAGATTCGCAATT 1533  
QY 501 AlaArgAlaLeuValArgAsnProLysileLeuileLeuAspGluAlaThrSerAlaLeu 520  
DB 1534 GCTCTGTCCTTAGTTTCGAAACCCCAAGATTTCTGATTTTAGATGAGGCTAGCTCTGCCCTG 1593  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 540  
DB 1594 GATTTCAGAAAGCAAGTCAAGTGTTCAGCTGCTGAGAGGATACCCCGAGTATTCA 1653  
QY 541 Phe 541  
DB 1654 TTT 1656  
RESULT 2

AX339031 AX339031 3702 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 13 from Patent WO0194400.  
DEFINITION AX339031  
ACCESSION AX339031  
VERSION AX339031.1 GI:18129123  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Frank, M.H. and Sayegh, M.H.  
AUTHORS  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 13 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
Location/Qualifiers  
source 1..3702  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1098 a 708 c 856 g 1039 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,04e-210 Length: 3702  
Score: 2697.00 Matches: 534  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.47% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-7 (1-541) x AX339031 (1-3702)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 34 ATGATCCTGGGTATACCTGGATCCTGGTCAATGGAGCCCTGCTTCTTAATGGCCACTG 93  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 94 GTTTTAGGAGAAATGATGATTAACCTTATTAGTGGATGCTAGTCCAAACTAACACATAC 153  
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
Db 154 TCTTTCTTCAGGTGACCTCTGATATATGTTGGAAATAGGTGTTGCTGCCTTGATTTTGGT 213  
Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
Db 214 TACATACAGATTTCTTGTGGATTATAACTGTCAGCAGCAGACCAAGAGGATTCGAA 273  
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100  
Db 274 CAGTTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTGATAGCTGTGACATCGGT 333  
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 334 GAACCTTAACACTCCCATGACAGACATTTGACAAATCAGTGTATGAGATTAAGATT 393  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 394 GCTCTGTGTTTCARAAACATGCTACTTTTTCATTTGGCTGGCAGTTGGTTGGTGAAG 453  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 454 GGCTGGAAACTACCTTAGTACTCTATCCAGTCTCTCTCTTATATAGCTTCAGCGGCA 513  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 574 GGGGCTGTGGCAGAGAAGCTTGTGTCATCAATCCGACAGTCATAGCTTTTAGGCCCCAG

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 634 GAGAAAGAACTTCAAAGGTCTTTCTTTAAATATAACAAGATATGCTTGGTGTATTTT 693  
Qy 221 ProGlnTTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 694 CCCAGTGGCTACTAAGTTGTCTGTTNTTTGTAGGTATACACAGATCTCAAGAT 753  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 754 GCAAGAGATTTTGGCATAAAGGACTATAGCTTCAAAGTCTCTCTGGTGTGTGTAC 813  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 814 TTCTTTATGAATGGAAACCTATGAGCTTGTCTTTTGGTATGGAACTCTCTTGATCTTAAT 873  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 874 GGAGAACCTGGATATACCATCGGACTGTCTTGTCTGTTTCTTTAGTGTATTCATAGC 933  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 934 AGTTATTGCTATGGAGCAGCTCCTCACTTTGAAACCTTCGCAATAGCCGAGGAGCT 993  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 994 GCCTTTTCATATTTTCAGGTTTATGATAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053  
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1054 GGATATAAACCTGAATCCATAGAGAACTGTGGAAATTTAAATAATGTTCTTCTTCAATTAT 1113  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1114 CCATCAAGACCATCTATCAAGATCTGAAAGGTCTGAATCTCAGAATTAAGTCGAGAG 1173  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1174 ACAGTCGCCCTTGGCTGCTCAATGCGCAGTGGGAGAGTACGGTAGTCCAGCTTCTGCAG 1233  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1234 AGTTATATGATCCGGATGATGCTTTATCATGGTGGATGAGATGACATCAGAGCTTTA 1293  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1294 AATGTCGCGCATATCGACACCATATTCGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1353  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1414 AGAGCAGCAAGGAAAGCAATGCGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1474 ACATTTGTTAGGGGAAAAAGGAGCTCAATGAGTGGAGGCGCAGAAACAGAGGATCGCAAT 1533  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1534 GCTCGTGCCTTAGTTGCAAAACCCCAAGATTTCTGATTTTAGATGAGGCTAGCTCTGCCCTG 1593  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1594 GATTCAAGAAAGCAAGTCAGCTGTTCAAGCTGCCTCGAGAG 1635  
RESULT 3  
AX339034 AX339034 1940 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 16 from Patent WO0194400.  
DEFINITION AX339034  
ACCESSION AX339034

VERSION AX339034.1 GI:18129126  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 16 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
Location/Qualifiers  
1..1940  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 530 a 394 c 448 g 568 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,45e-200 Length: 1940  
Score: 2568.50 Matches: 514  
Percent Similarity: 95.01% Conservative: 0  
Best Local Similarity: 95.01% Mismatches: 0  
Query Match: 93.78% Indels: 27  
DB: 6 Gaps: 1  
US-09-873-409-7 (1-541) x AX339034 (1-1940)  
QY 1 MetLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 34 ATGATCTGGGTATCTAGTGGATCACTGGTCAATGGAGCTTCCTCTTAATGCCACTG 93  
QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 94 GTTTTAGGAGAAATGAGTGATACCTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153  
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLeuLeuPheGly 60  
DB 154 TCITTCCTCAGTGTACCTGTATTATGTGGATAGTGTGTCTGCTCTGATTTTGGT 213  
QY 61 TyrIleGlnLeuSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgIleArgLys 80  
DB 214 TACATACAGATTTCTCTGTGGATTATTAACCTGCAGCAGCAGACCAAGAGGATTCGAAAA 273  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
DB 274 CAGTTTTTTCATTTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333  
QY 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 334 GAACCTTAACACTCCGATGACAGACATTTGACAAATCAGTATGGTATTGGAGATAGATT 393  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTCGCTGGCAGTTGGTTGGTGAAG 453  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 454 GGCTGGAAACTCACCTAGTAGTACTATCCACAGTCTCCTCTTATAATGGCTTCAGCGGCA 513  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 514 GCATGTTCTTAGATGGTCACTCATCTATTGACCNAGTAAAGAAATTAAGTGCCTATTCCAAAGCT 573  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 574 GGGCTGTGGCAGAGAAGTCTTGTCTCATCAATCCGAACAGTCAATAGCTTTAGGGCCAG 633  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 634 GAGAAAGAACTTCAA----- 648

QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 649 -----AGGTATACACAGAATCTCAAGAT 672  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 673 GCAAGAGATTTGGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTGTAC 732  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
DB 733 TTTCTTTATGAATGAACCTATGGACTGTCTTTTGGTATGGAACCTCTTGTATCTTAAT 792  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 793 GGAGAACCTGGATATACCATCGGACTGTCTTGTCTTTCTTTAGTGTAAATCCATAGC 852  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAACCTTCGCAATAGCCGAGAGCT 912  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 913 GCCTTTTCATATTTCAGGTTATTGATAGAAACCCAGATATAGATAACTTTTCCAGCT 972  
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 973 GGATATAAACCTGAAATCCATAGAGGAACTGTGGAATTTAAATAATCTTTCTTCAATTAT 1032  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTCAATCTCAGAAATTAAGTCTGGAG 1092  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1093 ACAGTCCCTCTGCTCGTCTCAATGTCAGTGGAGAGAGTACGCTAGTCCAGCTTTCGAG 1152  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspLysAsnAspIleArgAlaLeu 420  
DB 1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1213 AATGTGCGGCAATTATCGAGCCATATTTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1272  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
DB 1273 ACCACCATCAGTAACAAATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAG 1332  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1333 AGAGCAGCAAGGAGCAAAATGCTATGATTTTATCATGGAGTTTCTTAATAATTTAAT 1392  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1393 ACATTGTGAGGGGAAAAAGGAGCTCAAAATCAGTGGAGGGCAGAAACAGAGGATTCGCAAT 1452  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 1453 GCTCTGTCCTTAGTTCGAACCCCAAGATTCTGATTTTAGTAGGCTAGCTCTGCCCTG 1512  
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 540  
DB 1513 GATTTCAGAAAGCAAGTCAGCTGTTTCAAGCTGCACTGGAGAGGATACCCCAAGGATTTCA 1572  
QY 541 Phe 541  
DB 1573 TTT 1575  
RESULT 4  
AX339032 AX339032 3621 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 14 from Patent WO0194400.  
DEFINITION  
ACCESSION AX339032

```

VERSION      AX339032.1  GI:18129124
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        Frank, M.H. and Sayegh, M.H.
JOURNAL      A gene encoding a multidrug resistance human p-glycoprotein
              homologous to chromosome 7p15-21 and uses thereof
              Patent: WO 0194400-A 14 13-DEC-2001;
              THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES     Location/Qualifiers
              i. 3621
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
BASE COUNT   1081 a 696 c 842 g 1002 t
ORIGIN
Alignment Scores:
Pred. No.:    5.89e-197      Length:    3621
Score:        2528.50       Matches:  507
Percent Similarity: 94.94%   Conservative: 0
Best Local Similarity: 94.94% Mismatches: 0
Query Match:   92.31%      Indels:    27
DB:            6           Gaps:     1

US-09-873-409-7 (1-541) x AX339032 (1-3621)
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 34 ATGATCTGGGTATACCTGGATCACTGGTCAATGGAGCCCTGCTTCTTAAATGCCACTG 93
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
DB 94 GTTTTAGGAGAAATGAGTATACCTTATTAGTGGATGCTAGTCCAACTAACACATAC 153
QY 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePheGly 60
DB 154 TCTTTCTTCAGGTGACCTGTATTATGTTGAATAGGTGCTGCTTGAATTTTGGT 213
QY 61 TyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrIleArgGly 80
DB 214 TACATCAGATTTCTTGTGGATATTAACATGCAGCAGCAGCAGAGAGATTCGAAA 273
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100
DB 274 CAGTTTTTTCATTCAGTTTTGGCAGCAGACATCGCTGGTTGATAGCTGTGATCGGT 333
QY 101 GluLeuAsnThrArgMetThrAspIleAspIleSerAspGlyIleGlyAspIleGly 120
DB 334 GAACCTTAACACTCGCATGCAGACATTTGACAAATCAGTCATGTTGAGATAAGATT 393
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 394 GCTCTGTGTTTCAAAACATGTCTACTTTTTCGATTTGGCCCTGGCAGTTGTTGGTGAAG 453
QY 141 GlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 454 GGGTGAACCTCACCTAGTACCTCTATCCACGCTCTCTTATTAATGGCTTCAGCGCA 513
QY 161 AlaCysSerArgMetValIleSerLeuThrSerIleSerGlyLeuSerAlaTyrSerLysAla 180
DB 514 GCATGTTCTAGGATGGTCATCTCATTGACCAAGTAAAGGAATTAAGTGCCTATTCCAAAGCT 573
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 574 GGGGCTGTGGCAGAGAAGTCTTGTATCAATCCGACAGCTATAGCTTTTAGGGCCAC 633
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
DB 634 GAGAAAGAACTTCA----- 648

```

## RESULT 5

AX478104

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AX478104

Sequence 34 from Patent WO0240541.

AX478104

AX478104.1 GI:22217064

human.

Homo sapiens

3699 bp

DNA

linear

PAT 12-AUG-2002



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

# REFERENCE AUTHORS

Tang, Y. T., Yue, H., Nguyen, D. B., Hafalia, A. J., Elliott, V. S., Lu, Y., Walla, N. K., Yao, M. G., Baughn, M. R., Gandhi, A. R., Ding, L., Sanjanwala, M., Rankumar, J., Arvizu, C., Gietzen, K. J., Lal, P. G., Azimzai, Y., Khan, F. A., Thangavelu, K., Thornton, M., Lu, D. A., Tribouley, C. M., Warren, B. A., Ison, C. H., Das, D., Raumann, B. E., Policky, J. L., and Kearney, L.

Transporters and ion channels

Patent: WO 0240541-A 34 23-MAY-2002;

Incyte Genomics, Inc. (US)

## FEATURES

Location/Qualifiers

source

1..3699

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/notes="Incyte ID No: 7472030CB1"

BASE COUNT 1116 a 707 c 860 g 1016 t

## ORIGIN

Alignment Scores:  
Pred. No.: 1..81e-189 Length: 3699  
Score: 2437.00 Matches: 495  
Percent Similarity: 90.71% Conservative: 3  
Best Local Similarity: 90.16% Mismatches: 9  
Query Match: 88.97% Indels: 42  
DB: 6 Gaps: 3

US-09-873-409-7 (1-541) x AX478104 (1-3699)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 145 ATGATCTGGGTACTGACATCACTGTTCAATGGAGCTGCTTCTTAAATGCCACTG 204  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 205 TGTATAGAGAAATGAGTGAATACCTTATTAGTGGATGCTAGTCCACACATAACA 264  
Qy 41 SerPhePheArg-----LeuThr 46  
Db 265 AATTATCAGAACTGACTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324  
Qy 47 LeuTyrTrpValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
Db 325 CTGATTATGTTGGAAATAGGTGTTGCTGCTTGAATTTGGTTACATACAGATTTCTCTG 384  
Qy 67 TrpIlelethrAlaAlaArgGlnThrLysArgIleArgLysGlnPheHisSerVal 86  
Db 385 TGGATTATAACTGCAGCAGCAGACCAAGAGGATTCGAAACACAGTTTTTTTCATTCA 444  
Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
Db 445 TTGGCAGAGGACATCGGCTGTTGATAGCTGTGACATCGGTGAACCTTAACCTCGCATG 504  
Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
Db 505 ACAGATGACATTGACAAATCAGTGATGTTGGAGATAAGATTGCTCTGTTTCAA 564  
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145  
Db 565 AACATGTCTACTTTTTCATTTGGCTGCGAGTTGGTTTGGTGAAGGGCTGGAACCTCAC 624  
Qy 146 LeuValThrLeuSerThrSerProleuIleMetAlaSerAlaAlaCysSerArgMet 165  
Db 625 CTAGTGACTCTATCCAGCTCTCTTATAATGAGCTTCAGCGGAGCATGTTCTAGGATG 684  
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyValAlaValAlaGlu 185  
Db 685 GTCATCTATTGACAGTAAGGAATTAGTGCTTATCCAAAGCTGGGGCTGTGGCAGAA 744  
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205  
Db 745 GAAGTCTTGTCTATCAATCCGAACAGTCACTAGCTTTTAGGGCCCGAGGAGAAAGACTTCAA 804

Qy 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225  
Db 804 ----- 804  
Qy 226 SerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245  
Db 805 -----AGGTATACACAGATCTCAAGATGCAAGAGGATTTTGGC 843  
Qy 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265  
Db 844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTCTGTACTTCTTTATGAATGGA 903  
Qy 266 ThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285  
Db 904 ACCTATGGACTTCTTTTGGTATGGAACCTCTCTGATCTTAATGGAGAACCTGGATAT 963  
Qy 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305  
Db 964 ACCATCGGACTGTTCTTCTGCTGTTTCTTTAGTGTATATCCATAGTATGATTCATTGGA 1023  
Qy 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325  
Db 1024 GCAGCAGTCCCTCACCTTTGAAACCTTCGCAATAGCCCGAGAGCTGCTTTCAATTTTC 1083  
Qy 326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345  
Db 1084 CAGTTATTGATAAGAAACCCAGTATAGTAACTTTTCCACAGCTGATATAACTGAA 1143  
Qy 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365  
Db 1144 TCCATGAAGAACTGTGGAATTTAAATAATGTTTCTTTCAATTTATCCATCAAGGCCATCT 1203  
Qy 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385  
Db 1204 ATCAAGATTTCTGAAAGGTCTGGAATCTCGAATTAAGTCTGGAGAGACAGTCGCTTGGTC 1263  
Qy 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405  
Db 1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGTTTATATGATCCG 1323  
Qy 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425  
Db 1324 GATGATGGCTTTTATCATGTGTGATGAGAATGACATCAGAGCTTTAAATGTGCGCATAT 1383  
Qy 426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445  
Db 1384 CGAGACCATATTGGAGTGGTGTAGTCAAGAGCTGTTTGTTCGGAGCACCATCAGTAAC 1443  
Qy 446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465  
Db 1444 AATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAGGAA 1503  
Qy 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485  
Db 1504 GCAATTCGTATGATTTTATCATGGAGTTCCTTAATAAATTAATACATTTGGTAGGGGAA 1563  
Qy 486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505  
Db 1564 AAGAGACTCAATAGTGGAGGCGAGAAACAGAGGATCGCAATTTGCTCGTCTTAGTT 1623  
Qy 506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525  
Db 1624 CGAAACCCCAAGATTTCTGATTTTAGATGAGCTACGCTGCTGCTGGATTCAGAAAGCAAG 1683  
Qy 526 SerAlaValGlnAlaAlaLeuGluLys 534  
Db 1684 TCAGCTGTTCAAGCTGCACTGGAGAAG 1710  
RESULT 6  
AX339030 3177 bp DNA linear PAT 09-JAN-2002  
LOCUS  
DEFINITION Sequence 12 from Patent WO0194400.

ACCESSION AX339030  
VERSION AX339030.1 GI:18129122  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Frank, M.H. and Sayegh, M.H.  
AUTHORS A gene encoding a multidrug resistance human p-glycoprotein  
TITLE homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 12 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES Location/Qualifiers  
source 1..3177  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 970 a 501 c 735 g 870 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,58e-143 Length: 3177  
Score: 1868.00 Matches: 370  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 68.20% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-7 (1-541) x AX339030 (1-3177)  
QY 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184  
Db 1 ATGGTCATCTCATTACCAAGTAAGTAAAGTGCCTATTCCAAAGCTGGGCTGGCA 60  
QY 185 GluGluValLeuSerSerIleAaGThrValIleAlaPheAaGlaGlnGluLysGluLeu 204  
Db 61 GAAGAGTCTTGTTCATCAATCCGAACAGTCATAGCCCTTTAGGGCCCGAGGAAGAAT 120  
QY 205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224  
Db 121 CAAGGTCTTCTTTAAATATAACAGATATGCTTGTGTTTATTTCCCGAGTGGCTA 180  
QY 225 LeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244  
Db 181 CTAAGTGTGTTCTGTTNTTGTAAAGTATACACAGAACTCTCAAGATGCAAAAGGATTT 240  
QY 245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264  
Db 241 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTTTGGTGTGTACTTCTTTATGAAT 300  
QY 265 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284  
Db 301 GGAACTATGGACTTGTCTTTGGTATGGAACTCTCTGATCTTATATGAGAACTTGA 360  
QY 285 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIle 304  
Db 361 TATACCATCGGACTGTCTTCTGTTCTTTAGTGTATCCATACAGCATTATTCGATT 420  
QY 305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324  
Db 421 GGAGCAGCAGCTCCCTCATTGTAAACCTTCGCAATAGCCCGAGGAGTGCCTTCATATT 480  
QY 325 PheGlnValIleAspLysValProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344  
Db 481 TTCAGGTTATTGATAGAAACCCAGTATAGATAAATTTTCCACAGCTGGATATAACCT 540  
QY 345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364  
Db 541 GAATCCATAGAAGAACTGTGGAAATTAAATAATGTTCTTTCATATTATCCATCAAGACCA 600  
QY 365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384  
Db 601 TCTATCAAGATTCTGAAGGCTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCTTG 660

QY 385 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404  
Db 661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCGCAGAGTTATATGAT 720  
QY 405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424  
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAATGATCATCAGAGCTTTAAATGTGGCGCAT 780  
QY 425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444  
Db 781 TATCGAGACCATATTGGAGTGTAGTCAAGAGCCTGTTTGTGGGACCACCATCAGT 840  
QY 445 AsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 464  
Db 841 AACAATATCAAGTATCGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAGCAAG 900  
QY 465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484  
Db 901 GAAGCAATCGGTATGATTTTATCATGGAGTTTCTTAATAATTTTATCATTTGGTAGG 960  
QY 485 GluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504  
Db 961 GAAAAGGAGCTCAATGAGTGGAGGGCAGAAACAGAGATCGCAATTCCTGTCCTTA 1020  
QY 505 ValArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524  
Db 1021 GTTCGAAACCCCAAGATCTGATTTTAGATGAGGCTACGTCGCTCGCTTGCAGAAAGC 1080  
QY 525 LysSerAlaValGlnAlaLeuGluLys 534  
Db 1081 AAGTCAGCTGTTCAAGCTGCCTGGAGAAG 1110  
RESULT 7  
AX3322791 4189 bp DNA linear PAT 07-JAN-2002  
LOCUS AX3322791 Sequence 5 from Patent WO0192877.  
DEFINITION AX3322791  
ACCESSION AX3322791  
VERSION AX3322791.1 GI:18093768  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
REFERENCE Sorrentino, B. and Schuetz, J.  
AUTHORS Method of identifying and/or isolating stem cells  
TITLE Patent: WO 0192877-A 5 06-DEC-2001;  
JOURNAL ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)  
FEATURES Location/Qualifiers  
source 1..4189  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
BASE COUNT 1204 a 875 c 1028 g 1082 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,78e-118 Length: 4189  
Score: 1566.50 Matches: 307  
Percent Similarity: 72.63% Conservatives: 99  
Best Local Similarity: 54.92% Mismatches: 94  
Query Match: 57.19% Indels: 59  
DB: 6 Gaps: 6  
US-09-873-409-7 (1-541) x AX3322791 (1-4189)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 148 ATGATTCTGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCTCTCTGTGCTG 207  
QY 21 ValLeuGlyCysLeuMetSerAspAsnLeuIle----- 30  
Db 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTAGTATTCTGCAAGCAT 267

```
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAGTGGACCAACAGTACTCTGATCATCAGCAACAGCATCTCGGAGGAGAG 327
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGCCATATACGCCTAC-----TATTACACCGGATTTGGTGTGTTGTTG 372
QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLys 76
Db 373 CTCATAGTTCCTACATCCAGGTTTCACTTTGGTCTCGGAGCTGGAACACAGATACAC 432
QY 77 ArgIleArgLysGlnPhePheHisValLeuAlaGlnAspIleGlyTyrPheAspSer 96
Db 433 AAGATTAGGAGAGATTTTCCATGCTATAATGAATCAGGAGATAGCTGGTTGATGTG 492
QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 493 CATGATGTTGGGAGCTCAACACCCGGCTCACAGATGATGCTCCAAAATTAATGACGGA 552
QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 553 ATTGGTGACAAAATTTGGGATGTTTTTTCAGTCCATAACCAATTTTAGCGGTTTTATC 612
QY 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 613 ATAGATTATTAAGTGTGGAGCTTAACCTTGTCATTTTGGCTGTGACCCCTTTATT 672
QY 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 673 GGATTGTCATCTGCTTTGGGCAAGGATTTGATCTTCACTTACTAATAAGGAATCCAG 732
QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 733 GCTTATGCAAAAGCTGGAGCAGTGTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTG 792
QY 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db 793 GCCTTTGGAGACACAGAGAACTTGAA----- 822
QY 216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db 823 -----AGGTACAAT 831
QY 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db 832 AAAAAATTTAGAAAGAGCTAAATAATTTGGCATAAAGAAAGCTATCACAGCCAGCATTTGC 891
QY 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThr 275
Db 892 ATAGCATTCGCTACCTGTTGGTCTATGTCATCATATGCATCGGCATCTCTGGTATGGACA 951
QY 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db 952 TCCTTGGTCTCTCAATGAA-----TATTCTATTGGAGAGTCTTACTGCTCTCTTC 1005
QY 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db 1006 TCTATTTTGTTGGGACTTTTAGTATTGGACACTTGGCCCCCAACATAGAACCTTTGCA 1065
QY 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db 1066 AACGACGAGGGGAGCGCTTTGAATCTTCAAGATAATTGATACGAGCCAGCAATTTGAC 1125
QY 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsn 355
Db 1126 AGCTTCTCAACAAAGGGCTACAAACACAGACAGTAGTATAATGGGAAACTTAGAGTTTAAAAAT 1185
QY 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db 1186 GTTCATCTTCAACTACCCATCAGAAAGCAAGTTCAGATCTTTGAAGGGGCTCAATCTGAAG 1245
```

```
QY 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db 1246 GTCAAGAGCGGACAGACAGCGTGGCTTTGGTTGGCAACAGTGGCTGTGGAATAACCAACT 1305
QY 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db 1306 GTCACGCTGATCAGAGGCTCTACAGCCCTCGAGGGCGTGTGTATCGACGGACAA 1365
QY 416 AspIleArgAlaAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db 1366 GACATCAGAACCATCATGTGAGTATCTCAGGAGAGATCATTTGTTGGTGTGAGTACAGAA 1425
QY 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr 455
Db 1426 CCTGTGCTGTTGCCACCATCGCCGAGCAATTCGCTATGCGCCGAGAAGATGTCAAC 1485
QY 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db 1486 ATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCCCTATCATCATGAACATG 1545
QY 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLys 495
Db 1546 CCCACCAATTTGACACCTCTGTTGGTGAGAGAGGGCGCAGCTGATGGGGACAGAAA 1605
QY 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
Db 1606 CAGAGAATGCCATTGCCGGGCGCTGGTCCGCAATCCCAAGATCCTTTTGTGGACGAG 1665
QY 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGTGAGGCGGCCTGATGAAG 1722

RESULT 8
MUSMUR  Mouse multidrug resistant protein mRNA linear ROD 04-OCT-1995
LOCUS  M14757
DEFINITION  M14757.1 GI:199100
ACCESSION  M14757
VERSION  M14757.1
KEYWORDS  membrane-associated glycoprotein; multidrug resistance protein.
SOURCE  Mus musculus (strain BALB/c, sub_species domesticus) (clone: lambda
        DR11) cDNA to mRNA.
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 4298)
AUTHORS  Gros, P., Croop, J. and Housman, D.
TITLE  Mammalian multidrug resistance gene: complete cDNA sequence
        indicates strong homology to bacterial transport proteins
JOURNAL  Cell 47 (3), 371-380 (1996)
MEDLINE  87028229
PUBMED  3768958
COMMENT  The mdr gene family includes two or more related but distinct
        cellular genes.
        There are 3 possible start codons besides the one given in the
        Features table, at positions 161-163, 221-223 and 257-259. If the
        latter two start codons are used as initiator the first hydrophobic
        domain could possibly serve as a signal sequence.
FEATURES             Location/Qualifiers
     source           1..4298
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /clone="lambda DR11"
                     /cell_type="pre-B cell"
                     <1..4298
                     /product="mdr mRNA"
                     110..3940
                     /codon_start=1
                     /product="multidrug resistance protein"
                     /protein_id="AAA79005.1"
                     /db_xref="GI:387426"
                     /translation="MEFEENLGRADKNFSKMGKSKKEKKPAVGVFQMFYADW
```

LDKMLILGTLAAIHTGILLPLMLVFGNWDSTFKAEASILPSITNOSGPNSTLIIS  
 NSLSEEMAIYAYITGIGAGVLIIVAYIQVSLWCLAAGRIHKKIQKFFHAIWNQIEIG  
 WFDVHVGELNRLTDDSKINDGIGDKIMFQSIITPLAGFIIGFISGHWKLTIVIL  
 NVNPLIGLSSALAKVLTSTNKEIYAKAGAAVEALAAITVIAFGQQQKLEERY  
 AKNLEAKNVGKKAITASISIGIAYLVYASALAFYGTSLVLSNYSYISGIVLTVF  
 FSILAGTSGHGLANIEAFANAGAAPEIFKIIDNEPSIDSFSTKVYKPSIDMGNLE  
 FKVNHPNPSREVOILKGLNLKVKSGQTVVALVGNCGCKSTTVQLMQRLYDPLSGVV  
 SIDQDITRTNRYLREIIGVVSQBPVLFATTIAENIRYREDVDMDEIEKAKGANA  
 YDFIMKLPQFTLVGERAQSLGGQKQRIATARALVRNPKILLIDDEATSLDTRSEA  
 VVQAAIDKAREGRTTIVIAHRLSTVRNADVIAGFDGCVIVEQGNHDELMREKGIYFKL  
 VMTQTRGNEIPBNNAYSQSDTASELTSSEKSPILRRSYIYRSHRKODJDERLSM  
 KEAVDEDPVLSFMRILANLSEMPVLLGVLCVINGCIPQVFAIVSRJVFPSRD  
 DHEHTRQKCNFSIFLLVNGLISFVYFFQSTFGKAGEILITKVRVNFVMSMLRQD  
 ISWFDHKNSTGLSTRILASDASSVKGAMGARLAVVTONVNLGTGTVILSVYGMOLT  
 LLLVITLPIVLVIGLIEIMKLLSGQALKQKLEISGKIATEAIEINFRITVSITRQKF  
 ETMYAQSVQVRYNNAKMAHVEGTFPSFTQAMMYFSYAACFPFGAYLVAQQUMTFENV  
 MLVFSAVFGMAAGNDSFADYAKAKVSASHIRIIRIKTEPIDSYEGELKPYLLE  
 GNVKNGVOPNPTRNPIVLOGLSLEVKKGTALVGVSSGGCKSTVVQVLLERFYDPM  
 AGSVFLDQKEIKQLNQMWRALHGLIVSQBPILFDCSIAENIAYGNSRAVSEEBIVRA  
 AKEANIHQFIDSLPKINTRVGDGKGTQLSGQKQRIATARALVRNPKILLIDDEATSL  
 DTSEKVVQVQALDKAREGRTTIVIAHRLSTIQAHLSTIQADLVIVENGKVEHGHQOLLAQK  
 GYFSMVQAGAKRS"

BASE COUNT 1226 a 903 c 1059 g 1110 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,88e-118 Length: 4298  
 Score: 1566.50 Matches: 307  
 Percent Similarity: 72.63% Conservative: 99  
 Best Local Similarity: 54.92% Mismatches: 94  
 Query Match: 59 Indels: 59  
 DB: 10 Gaps: 6

US-09-873-409-7 (1-541) x MUSMDR (1-4298)

QY 1 MetileuGlyleleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 257 ATGATTCTGGGAACCTCTCGTGTCTATTATCATCGGAACATTACTTCCCTCTTGTGCTG 316  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuLe- 30  
 Db 317 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTATTCTGCCAAGCAATT 376  
 QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
 Db 377 ACTAATCAAGTGGACCCACAGTACTCTGATCATCAGCAACAGCAGCTCTGGAGGAAGAG 436  
 QY 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlylleGlyValAlaAla 56  
 Db 437 ATGCCCATATACGCCCTAC-----TATTACACCGGGATTGCTGCTGCTG 481  
 QY 57 LeuilePheGlyTyrileGlnileSerLeuThrileThrAlaAlaArgGlnThrLys 76  
 Db 482 CTATAGTTCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAACAGATACAC 541  
 QY 77 ArgileArgLysGlnPhePheHisSerValLeuAlaGlnAspilleGlyTyrPheAspSer 96  
 Db 542 AAGATTAGGCAAGATTTTCCATGCTATAATGAATCAGGAGATAGCTGCTGTTGATGTG 601  
 QY 97 CysAspilleGlyGluLeuAsnThrArgMetThr---AspileAspLysileSerAspGly 115  
 Db 602 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGCTCTCCAAAATAATAGCGGA 661  
 QY 116 IleGlyAspLysileAlaLeuLeuPheGlnAsnMetSerThrPheSerilleGlyLeuAla 135  
 Db 662 ATGGGTGACAAAATTTGGGATGTTTTTTCAGTCCATAACCAACATTTTATAGCCGGTTTTATC 721  
 QY 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuile 155  
 Db 722 ATAGATTATTAAGTGGTTGGAAGCTTAACCTGTGCTATTTTGGCTGTACAGCCCTCTTATT 781  
 QY 156 MetAlaSerAlaAlaAlaCysSerArgMetValileSerLeuThrSerLysGluLeuSer 175

Db 782 GGATTGTCATCTGCTTGTGGCAAGAGTATTGACTTCTATTACTAATAGGAACCTCCAG 841  
 QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerileArgThrValile 195  
 Db 842 GCTTATGCAAAAGCTGGAGCAGTGTCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 901  
 QY 196 AlaPheArgAlaGlnGlnLysGluLeuGlnArgSerPheLeuLeuAsnileThrArgTyr 215  
 Db 902 GCCTTTGGAGGACACAGAAAGAACTTGA----- 931  
 QY 216 AlaTyrPheTyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThr 235  
 Db 932 -----AGGTACAAAT 940  
 QY 236 GlnAsnLeuLysAspAlaLysAspPheGlyileLysArgThrileAlaSerLysValSer 255  
 Db 941 AAAAAATTTAGAAAGCTAAAAATGTTGGCATAAAGAAAGCTATCAGCAGCAGCAATTTCG 1000  
 QY 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThr 275  
 Db 1001 ATAGGCAATTGCCTTACCTGTTGTTCTATGCATCATATGCACTGGCATTTCTGGTATGGACA 1060  
 QY 276 SerLeuileLeuAsnGlyGluProGlyTyrThrileGlyThrValLeuAlaValPhePhe 295  
 Db 1061 TCTTGTGCTCTCAAAATGA-----TATTCTATTGGAGAGTGTCTTACTGCTCTTCTTC 1114  
 QY 296 SerValileHisSerSerTyrCysilleGlyAlaAlaValProHisPheGluThrPheAla 315  
 Db 1115 TCTATTGTTGTTGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAACCTTTTGA 1174  
 QY 316 IleAlaArgGlyAlaAlaPheHisilePheGlnValileAspLysLysProSerileAsp 335  
 Db 1175 AACGCACGAGGGGCGACCTTTGAAATCTTCAAGATTAATGATAACGAGCCCAACATTGAC 1234  
 QY 336 AsnPheSerThrAlaGlyTyrLysProGluSerileGluGlyThrValGluPheLysAsn 355  
 Db 1235 AGCTTCTCAACAAAGGCTTACAAACAGCAGTATTAATGGAAACTTAGAGTTTAAAT 1294  
 QY 356 ValSerPheAsnTyrProSerArgProSerileLysileLeuLysGlyLeuAsnLeuArg 375  
 Db 1295 GTTCACTTCACTTACCCATCGAGAGCGAAGTTTCAGATCTTGAAGGGCTCAATCTGAAG 1354  
 QY 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395  
 Db 1355 GTCAAGAGCGGACAGCAGCGTGGCTTGGTTGGCAACAGTGGCTGTGGAAAGAACCAACT 1414  
 QY 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheileMetValAspGluAsn 415  
 Db 1415 GTCCAGCTGATGCGAGGCTTACAGACCCCTCGAGGGGCTGTGTAGTATCGAGCGGACAA 1474  
 QY 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisileGlyValValSerGlnGlu 435  
 Db 1475 GACATCAGAACCATCAATGTAGGTATCTGAGGGAGATCATTTGTTGTTGAGTCAGGAA 1534  
 QY 436 ProValLeuPheGlyThrThrileSerAsnAsnileLysTyrGlyArgAspValThr 455  
 Db 1535 CCGTGTGCTGTTTCCACCAACGATCGCGAGAACATTCGCTATCGCGGAGAGATGTCAACC 1594  
 QY 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheileMetGluPhe 475  
 Db 1595 ATGGATGAGATGAGAAAGCTGTCAAGGAGAGCCAAATGCCCTATGACTTCATCATGAAGCT 1654  
 QY 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyValaGlnMetSerGlyGlyGlnLys 495  
 Db 1655 CCCACCAATTTGACACCTGTTGTTGTTGAGAGAGGGCGCAGCTGAGTGGGGACAGAAA 1714  
 QY 496 GlnArgileAlaAlaAlaArgAlaLeuValArgAsnProLysileLeuileLeuAspGlu 515  
 Db 1715 CAGAGAATCGCCATTCGCCGGGCGCTTGGTCCGCAATCCCAAGATCTCTTTTGTGGACGAG 1774  
 QY 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
 Db 1775 GCCACCTCGACCTGGATACAGAAAGTGAAGCTGTGTGTGCGAGCGCCGACCTGGATAG 1831

```

RESULT 9
AX105082
LOCUS AX105082 4279 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 26 from Patent WO0123540.
ACCESSION AX105082
VERSION AX105082.1 GI:13921232
KEYWORDS
SOURCE
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 4279)
AUTHORS
Stocker, P.J., Steinel-Crepi, D.T., Crepi, C.L., Reif, T.C. and
Patten, C.J.
TITLE
P-glycoproteins and uses thereof
JOURNAL
PATENT: WO 0123540-A 26 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
Location/Qualifiers
1..4279
/organism="Canis familiaris"
/db_xref="taxon:9615"
17..3862
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:13921233"
/translation="MDPEGRKGSAEKNFWMGKKKKKKKKKKKTVSTFAMFRYSN
WDLRLNMLVGTMAAIHGAALPLMWLVFGNMTDSFANAGISRNKTPFVINESITNNT
QHPINLHEEMTYAYYISGIGAGVLAAYIOVSFWCAGQILKIRKOFFHAINRQ
EIGFWDVHDVGLNLTDDVSKINGEDIGKIMFFQSIAATFTGIVGFRGWLTL
VILASIPVLGSLAAIKALISFTDKELIAYAKAGAVEBLAAIRTVIAFGOKKEL
ERYKNLEAKGIGIKATANTISIGNAFLIYASVALAFWYCTSLVASEVTIGVL
TVFVLIGAFSGQASPEAFANRGAAYEIPKIIDNKPISDTSKSGHKPDNKG
NLEFKNVHFSYRKEKILKLNKLVQSGQTVLVNCGSGKSTTVQLMQLDPTD
GMVCDQDRTINVRHLREITGVSOEPLVFATTAENIRYGRNVTWDEIBAKVE
ANAYDFTMKLPNKFDTLVGERGAOLSGGQKRIARALVNPVKIILLDEATSLDTE
SEAVOVALDKARKRTTIVIAHRLSTVRNADVIAGDGVIVKEGNHDELMEKGIY
FKLVTMOTRNEIPELVNATGESSEDALEMSPKDSGSLIKERSTERSHAPQGDOR
KLGTKEDLNENLEPNRILKLNSTENYFVVGIFCAINGGLQAPAFSIFRIIGI
FRDEDEFTKQNSNMFVSLFVLGIISFIFLFGFTFGKAGEILTKRLRYNVFRSM
LRQDSWDFKQNTALATRLANDAAQVKAIGSLAVITONAMLTGIIISLYG
WQTLALLAIVPIAIAVGVMKMLSGQALKDKKELEGAKIATEIENFRIVLSLR
EOKFEVYAOQLGVYNSLRKAHIFGVSPSITQMMYFSYACGFRGAVLVANEFN
FQDVLVFSIAIVGMAVQVSSFPADYAKAKVAAHVIMIEKSLPIDSYPHGLKP
NTLEGNVTFNEVFNVPTRDIPVIGLSLEVKKQTLALVSGSGCGCTVQQLLERF
YDPLAGSVLIDKEIKHLNWLRAHLGIYSQEPILFDQSIENIAYGDNRSVSHDEE
IVQAAKEANTHHFETLPEKYNTRVGDGKQTLGGQKQRIARALVROPHILLDEA
TSALDTESEKVVQEALDKAREGRTCIIVIAHRLSTIQNADLIVVFQNGKVEHGHQOL
LAQKGIYFSMVSVQAGAKR"
BASE COUNT 1296 a 833 c 1009 g 1141 t
ORIGIN
Alignment Scores:
Pred. No.: 7,32e-118 Length: 4279
Score: 1561.50 Matches: 304
Percent Similarity: 72.12% Conservative: 97
Best Local Similarity: 54.68% Mismatches: 104
Query Match: 57.01% Indels: 51
DB: 6 Gaps: 5
US-09-873-409-7 (1-541) x AX105082 (1-4279)
Qy 1 MetLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAAATGGCTGCCATCATCCATGAGCTGCACCTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGGAAACATGACAGATAGCTTTGCAAAATGAGAGAAATTCAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44

```

```

Db 287 CCAGTTATAATTAATGAAAGTATTATCAGAACATATACAAACATTTTCATCAACCACTTGGAG 346
Qy 45 -----LeuThrLeuTyrValGlyIleGlyValAlaLeuIlePhe 59
Db 347 GAGAAATGACACAGCTATGCTATTATTACAGTGGGATCGTGCTGGCGTGTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleAspArg 79
Db 407 GCTTACATCCAGGTTTCATCTGGTGGCAGCAGGAGAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATCGACGAGAGATGGCTGGTGGTGGTGGTGGTGGT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAsp 118
Db 527 GGGAGCTTAACACCGGCTCACAGAGATGCTCTCAAAATCAATGAAGAAATGGCGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAATTTGGAATGTTCTTCAATCAATAGCAACATTTTACCGGTTTATAGTGGGTTT 646
Qy 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTGGAGCTTAACCTTGTGATTTGGCCATCAGCCCTGTCTTGGACTTTCA 706
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCGCCATCTGGGCAAGATATCTTCTTACTGATAAAGAACTCTTGGCCTATGCA 766
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGACAGTAGCTGAAGAAGCTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 826
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 827 GGACAAAGAAAGAACTTGAA----- 847
Qy 219 TyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeu 238
Db 848 -----AGGTACAAACAAAATTTA 865
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 866 GAAGAAGCTTAAGGAATTTGGATAAGAAAGCTATCACGGCCCAACATTTCTATTGGTGCC 925
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 926 GCTTTCTTATTGATCTATGTCATCATGCTCTGGCTTTCTGGTATGGGACCTCTCTGGTC 985
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 986 CTCTCCAGTGAA-----TATACTATTGGACAGTACTCATCTGCTCTCTTTCTGTATTA 1039
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATGGGGCTTTAGTATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTATGAAATCTTCAAGATAATTCACAAATTAACCAAGCATTTGACAGCTATTCG 1159
Qy 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAACAGATATTAATTAAGGAAATTTGAAATTTCAAAATTTGTCATTTC 1219
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCTCTTCGAAAGAAGTTAAGATCTTAAAGGCTCTCAACCTGAAGTTTCAGAGT 1279
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1280 GGGCAGACAGTGGGCTGGTGGGAAACAGTGGCTGGCGGAAGACGACGCGGTGAGCTG 1339

```









```
QY 119 LysileAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAAATTGGAATGTTCTTCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGTTT 646
QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTGGAGCTTAACCTTGTGATTTGGCCATCAGCCCTGTTCTGGACATTTC 706
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCGCCCATCTGGGCAAGAGATCTATCTTCTACTGATAAAGAACTCTTGGCCTATGCA 766
QY 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAAGCTGATGCTTTGGA 826
QY 199 AlsGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db 827 GGACMAAGAAAGAACTTGAA----- 847
QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu**PheValArgTyrThrGlnAsnLeu 238
Db 848 -----AGGTACACAAAAATTGA 855
QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db 866 GAAGAAGCTTAAGGAATTTGGATAAAGAAAGCTATCATCAGGCCAACATTTCTATTGGTGC 925
QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db 926 GCTTCTTATTGATCTATGCATATATGCTCTGGCTTTCTGGTATGGACCTCTCTGGTC 985
QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db 986 CTCTCCAGTGAA-----TATCTTATTGGCAAGTACTCATGTCCTCTCTTCTGTTATTA 1039
QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATTTGGGCTTTAGTATTGGACAGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGA 1099
QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTTATGAATCTTCAAGATAATTCACATAATTAACAAACCAAGCATTCAGCATTCG 1159
QY 339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAACCAGATAATTTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219
QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCTTCTCGAAAAGAAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTACAGT 1279
QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTCGCGGGAAGACGACGCGTGCAGCTG 1339
QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1340 ATGCAGAGGCTCTATGACCCCAAGTGGCATGCTGTTGTTATGATGACAGGACATTAGG 1399
QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1400 ACCATAAATGTAAGCATCTTCGGGAATATTCTGTTGTTGATGTCAGGAGCTGTGTTG 1459
QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db 1460 TTTGGCCACCAGCATAGCTGAAAAACATTCGTTATGCGCCGCAAAATGTCACCATGATGAG 1519
QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAAGCTGTTAAGGAAGCCAAATGCCTATGATTTTATCATGAACATACCTATAAA 1579
```

```
QY 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1580 TTTGACACTCTGTTGGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAGAAATC 1639
QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTGGGCCCTGTTGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699
QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1700 GCTCTGCACACTGAAGTGAAGCAGCTGGTTCAGTGGCCCTGGATAAG 1747

RESULT 12
LOCUS AX108656 4195 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123565.
ACCESSION AX108656
VERSION AX108656.1 GI:13923888
KEYWORDS crab-eating macaque.
SOURCE Macaca fascicularis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 4195)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
TITLE p-glycoproteins from macaca fascicularis and uses thereof
JOURNAL Patent: WO 0123565-A 3 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
source
1..4195
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
100..3951
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:13923889"
/translation="MDLEGDRNGAEKKNFKLNKSKKDKKERKPTVSFMSFRYSN
WLDKLYMVVGTAAIHHGAGLPMLMLVFGDMTDTFANAGNLGDLGALLFNNTSSNIT
DTVPWNLEEDMTRYAYYSYGICAGVLVAAYLOVSFWCLAGROGHKIRKQFTHAMR
QEIGFVDHVDGELNTRLTDVSKINEGDKIGMFQSMATFFTFGIVIAFGQKKE
LVILAIPVLGSLAAVWAKIISFTDKELLAYAKAGAAVEALAAIRTVAFGQKKE
LERYNKLAEAKRIGIKKAITANISIGAFLLIYASALAFWYGTTLVLSKEYSIGOV
LTVFSLVIGFSPVGOASPTAPANARGAAPEIKIIDNKPISIDSKSGHKPDNIK
GNLEFRNVHFSYPSRKEVKILGNLKVQSGOTVALVNSGCGKSTTVOLMORLYDPT
EGMWSVDCQDRTINVREIREIIGVVSQEPVLFTAIENIRYGRDVTMDSEKAVK
ENAYYDFIMKUPQKEDTLVBERGAQSGQKQRIARALVRNPKILLUDEATSAUDT
ESEAYVVALDKARKGRTTIVIAHRLSTVRNADVIAGDFDGVIVKGNHDELMKEGI
YFKLVTMQTAGNEIELENAADESKSEIDTLEMSHDSGSLIRKSTRSRVSQSGQD
RKLSTKEALDESIIPVSWFRIMKMLTEWPYFVGVFCAIINGQLQPAFAVIFSKIIG
IFTRNDDAETKRONSLFSLFLVLGIVSFITFLOQFTFGKAGAILTKRLAYMVPFRS
MLRDVSNFDDPKNTGALTTRLANDAAQVKAIGSRLAITQNTIANLGTGIIISLIY
GMLTLLLLLAIVPIIAAGVEMVMKLSGALKDKKELEGALEATEAENFRVIVSLM
QEQKEFHYDQSLQVPYRNSLRKAHIFGITTFQAMMYFSYAGCFRFGAYVAHSLM
SPEDQLLVFSAVFVGMAVQGVSSFPAYAKAVSAAHIMIIEKTPIDISYSTEGLK
PNTLGNVTENEWENYPTRLDI PVQLGLSLVKKGQTLALVNSGCGKSTTVQLLER
FYDLPLAGVLLDGKEIKOLNQWMLRAHLGIVSQEPILPDCSISENIAYGDSNRVVSQE
ELVRAAKENIAHFIESLPNKYSTRVGDKTQSGQKQRIARALVRQPHILLIDE
EATSALDTESEKRVQEALDKAREGRTICVIAHRLSTQNLADLIIVFQNGVRKRGHGHQ
LLAQGIYFSMVSVQAGAKRQ"
BASE COUNT 1230 a 801 c 1039 g 1125 t
ORIGIN

Alignment Scores:
Pred. No.: 1,678-117 Length: 4195
Score: 1557.00 Matches: 306
Percent Similarity: 72.17% Conservative: 96
Best Local Similarity: 54.94% Mismatches: 103
Query Match: 56.85% Indels: 52
DB: 6 Gaps: 5

US-09-873-409-7 (1-541) x AX108656 (1-4195)
```

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 250 ATGGTGGTGGAACTTTGGCTGCATCATCCATGGAGCTGGAGTTCCTCTCATGATCGT 309
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33
Db 310 GTGTTTGGAGACATGACGGTACCTTTGGCAAAATGCAGAAATTTAGGAGATTTAGGAGCT 369
QY 34 LeuValGlnThrAenThrTyrSer----- 41
Db 370 CTGTTGTTTAAACACATAATAGCAGTAATATCACTGATCAGTGCCTGCATGAATCTG 429
QY 42 -----PhePheArgLeuThrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db 430 GAGGAAGATATGACCATGATGCTTATATACAGTGGAAATGGTGTGGGTGCTGTT 489
QY 59 PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db 490 GCTGCTTACATTTAGGTTTCAATTTGGTGGCTGGCAGCTGGAGACAAATACACAAAT 549
QY 79 ArgGlyGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp 98
Db 550 AGAAAAAGATTTTTCAGTCTATTAATGCGCAGGAGATAGCTGGTTGATGTGCAGAT 609
QY 99 IleGlyGluLeuAenThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db 610 GTTGGGAGCTTAACACCGGCTTACAGATGATGCTCCAAAGATTAATGAAGGAATGGT 669
QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db 670 GACAAAATTTGGAATGTTCTTTTCACTCAATGGCAACATTTTCACTGGGTTTATAGTAGGA 729
QY 138 LeuValLysGlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db 730 TTACAGCTGGTGGAGCTGAACCTTGTGATTTGGCCATCAGTCTCTGTTTGGAGCTG 789
QY 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db 790 TCAGCTCAGCTGCGGCAAGATACCTCTTCATTTACTGATATAAGAACTCTTAGCTTAT 849
QY 178 SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db 850 GCAAAAGCTGGAGCAGTAGCTGAAGAGTCTTGGCAGCAATTAGAACTGTGATTCATTT 909
QY 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyr 217
Db 910 GGAGGACAAAAGAAAGAACTCGAA----- 933
QY 218 PheTyrPheProGlnTyrLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db 934 -----AGGTACACAAAAT 948
QY 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db 949 TTAGAAGAGCTAAAGAAATTTGGATAAGAAAGCTTATACAGCAATATTTCTATAGT 1008
QY 258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeu 277
Db 1009 GCTGCTTCTGCTTATCTATGATCATCTATGCTCTGGCTCTGGTATGGAGCACCTTG 1068
QY 278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db 1069 GTCTCTCAAGGAA-----TATCTATTGGACAGTACTCACTGATTTCTTTCTGTA 1122
QY 298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db 1123 TTAATTTGGGCTTTAGTGTGGACAGGCATCTCCAGCATTTGAAGCATTTGCAAAATGCA 1182
QY 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db 1183 AGAGGAGCAGCTTTTGAATCTTCAAGATAATTTGATAATTAAGCAAGTATTGACAGCTAT 1242
```

```
QY 338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db 1243 TCGAAGAGTGGGCACAAACCCAGATAATATTAAAGGAAATTTGGAATTCAGAAATTTTCAC 1302
QY 358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db 1303 TTCAGTTTACCATCTCGAAAGAAAGTAAAGATCTTGAAGGGCCTGAACCTGAAGGTGCG 1362
QY 378 SerGlyGlyThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db 1363 AGTGGGAGACGCTGGCCCTGGTTGGAACAGCGGCTGTGGAGAGACCAACCGTCCAG 1422
QY 398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db 1423 CTGATGCAGAGGCTTTATGACCCACAGAGGCGATGCTCAGTGTGATGACAGGATATT 1482
QY 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db 1483 AGGACCATAAACGTAAGGTTTCTACGGGAATCATCGGTGTGCTGAGTCAAGAACCTGTA 1542
QY 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGlu 457
Db 1543 TTGTTTCCACACAGATAGCTGAAACATTCGCTATGCTGTAAGATGTCACCATGGAT 1602
QY 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db 1603 GAGATTTGAGAAAGCTGTCAAGGAAGCAATGCCTATGACTTTATCATGAAACTGCCTCAG 1662
QY 478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db 1663 AAATTTGACACCTGTTGGAGAGAGAGGGGCCAGCTGAGTGTGGGCAAGACGAGG 1722
QY 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db 1723 ATCCGATTCGACGTGCCCTGGTTTCGCAACCCCAAGATCTCTCTGTCGAGAGGCCAG 1782
QY 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1783 TCAGCTTGGACACAGAAAGTGAAGCAGTGGTTTCAGTGTGCTCTGGATAAG 1833
RESULT 13
AF269224 4045 bp DNA linear SYN 04-JUL-2000
LOCUS Synthetic construct Canis familiaris his-tagged-multidrug
DEFINITION resistance glycoprotein gene, complete cds.
ACCESSION AF269224
VERSION AF269224.1 GI:8926216
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE Analysis of dog MDR1 p-glycoprotein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4045)
AUTHORS Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
DIRECT SUBMISSION
JOURNAL Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
Tournepieuville, BP3, Toulouse 31931, France
FEATURES
source
1. .4045
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="original mRNA isolated from Beagle dog colon"
101. .3964
/codon_start=1
/transl_table=11
/product="his-tagged-multidrug resistance glycoprotein
MDR1"
/protein_id="AAF81747.1"
/db_xref="GI:8926217"
/translation="MHFHHHHDPEGGRKGSAEKFNFKMGKSKQNEKKEKPTVTSTFA
```

MFYSNWLDRLLVLTGTMAAIIHGAAALPLMLVFCNMWTDSPANAGISRNKTFPPVLIN  
SITNQFHNHLEEMETIYYISGIGAGVLVAAIYQVSWFCLAGAQILKIRKQFF  
HAIKMOEIGFDVHDVGNLRLTDDVSKINEIGDKIMFFQSTQFTGFIYVGFTR  
GWLKLVTLATISVPLGLSAAIWKALISSTDKELLAYAKAGAVEVLAIAIRTVIAFG  
GOKLEERYKNLEBAKIGIKALITANI SIGAAFLIYASVALAFVGTGSLVLSSEY  
SIGOVLTVFFSLIGAFSIGOASPSIEAFANARGAAVEIFKIDNKPISDSVSKGHK  
PDNIKGNLEFNKVPHSYPSRKEVKILKGNLKVQSGQTVVALVGNSSCKSTVQLMQR  
LYPTDGMVCDIGDQIRTNVHRLREITGVVSOEPLVFAITTAENIRIGRENVMTDEI  
EKAVANAYDPIMLKPNKFTLVGERGAQLSGGQKQRIATARALVNPKNILLDEAT  
SALDTSEAVQVQALDKRGRITTVIAHRLSTVRNADVIAGDDGVIVKGNHDELM  
KEKGYFKLVMTQNGNIELENATGESKSDALEMPKQSGSLIKRSLRRSRIHA  
POQDRKLGTEDENENPVVPSFWRLKLNSTEWPFVVGIFCAIINGSLGLOPASIIF  
SRIGIFRDEDEPETERKQNSNMFSLFVLGIIISFITFLOQFTFGKAGEILTKRLRY  
MYFRSMLQDVSFWDDPKNTTALJTRLMNDAQVKGAGISLAVITQNIANLGTGII  
ISLYLWQQLTLLLLAIVITAIAGVEMKMSGQALKDKELEGAKTATEIANENFT  
VVSLEKQFVEFYAQSIVPVYRNSLRKAHIFGVFSITQAMMYFSYAGCFRFGAYLV  
ANERNFODVLVFAISIVFGAMVQVSSFPADYAKAKVASAHVIMI TEKPLIDISY  
PHGLKPNLTLEGNVTNEVFNYPTRPDIPVLQGLSLEVKGGOTLALVSGSGCKSTV  
QLLERFYDPLAGSLIDGKEIKHLNVQMLRAHLGVQSORPILFDCSIAENIAYGNSR  
VVSHEEIMQAAKEANIHFIEIETLPEKYNTRVGDGKQTLQSGGQKQRIARALVRPHI  
LLLDDEATSDALDTSEKVVQEAQDREKRETCIVIAHRLSTIQNADLI VVVFQNGVKYEH  
GTHOQLLAQKGIYFSMISVQAKR"

BASE COUNT 1171 a 850 c 986 g 1038 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,75e-117 Length: 4045  
Score: 1556.50 Matches: 303  
Percent Similarity: 72.12% Conservative: 98  
Best Local Similarity: 54.50% Mismatches: 104  
Query Match: 56.83% Indels: 51  
DB: 12 Gaps: 5

US-09-873-409-7 (1-541) x AF269224 (1-4045)

Qy 1 MetIleuGlyIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 269 ATGTGGTGGGCAATGGCTGCATCATCCATGGAGCTGCACCTCCTCATGATGCTG 328  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuSerGlyCysLeu----- 34  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 329 GTTTTGGAAACATGACAGTACTTGTGCAATGCAGGAATTTCAAGAAACAACTTTT 388  
Qy 35 -----ValGlnThrAsnThrTy-SerPhePheArg----- 44  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 389 CCAGTTATAATTAAGTATTACGAACAATACACAACTTTCATCAACCATCTGGAG 448  
Qy 45 -----LeuThrLeuTyTyTyValGlyIleValAlaAlaLeuIlePhe 59  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 449 GAGGAAATGACACGCTATGCTTATTATACAGTGGGATCGGTGCTGGCTGGCT 508  
Qy 60 GlyTyIleGlnIleSerLeuTPIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 509 GCTTATCATCAGGTTTCATTCCTGGTGGGCGAGGAGGACAGATATCAAAATTAGA 568  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 569 AAACAATTTTTCATGCTATCATCGGACGAGGAGATGGCTGGTTTGACGTGCATGAGT 628  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 629 GGGGAGCTTAACACCGGCTCACAGACGATGTCTCAAAATCAATGAAGAAATGGCGAC 688  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 689 AAAATGGGAATGTTCTTCAATCAATGACGAACATTTTCAACCGGTTTATAGTGGGTTT 748  
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 749 ACAGTGGTGGAGGTAACCTTGTGATTTGGCCATCAGCCCTGTTCTTGACATTTCA 808  
Qy 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySer 178  
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14  
AX322787  
LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002  
DEFINITION Sequence 1 from Patent WO0192877.  
ACCESSION AX322787  
VERSION AX322787.1 GI:18093766  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Sorrentino B. and Schuetz J.  
TITLE Method of identifying and/or isolating stem cells  
JOURNAL Patent: WO 0192877-A 1 08-DEC-2001;  
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)  
FEATURES  
source  
1. 3860  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1135 a -746 c 957 g 1022 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.99e-117 Length: 3860  
Score: 1555.50 Matches: 306  
Percent Similarity: 73.29% Conservative: 100  
Best Local Similarity: 55.23% Mismatches: 99  
Query Match: 56.79% Indels: 49  
DB: 6 Gaps: 6  
US-09-873-409-7 (1-541) x AX322787 (1-3860)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 151 ATGGTGGGAACTTTGGCTGCCATCATCCATGGGCTGGAGTCTCTCTCATGATCGT 210  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATCGAGAAATTTAGAAATCTGATGCA 270  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 271 AACATCAATAAGATGATATCAATCATACAGGTTCTTCATGAATCTGGAGAAAGAC 330  
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 331 ATGACCATATATGCTTATTATACAGTGAATTTGGTGTGGGGTGTGGTGTGGTCTTAC 390  
Qy 62 IleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleAtqLysGln 81  
Db 391 ATTCAGGTTTCATTTGGTGTGGCTGGAGCTGGAGAGCAATATACACAAATTTAGAAACAG 450  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 451 TTTTTCATGCTATAATGCGACAGAGATAGCTGTGTTGATGTCACGATTTGGGGAG 510  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 511 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTATTGTTGGTGCACAAAT 570  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 571 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCATCTGGGTTTATAGTAGGATTTACACGT 630  
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 631 GTTGGAGACTAACCCCTGTGATTTTGGCCATCAGTCTCTGTCTTGGAGTCTCAGCTGCT 690  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 691 GTCCTGGGCAAGAGACTATCTTCTTATTAATGATAAGAACTCTTTAGCGTATGCAAAAGCT 750

Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 751 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATTGCAATTGGAGACAA 810  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 811 AAGAAAGAACTTGAA----- 825  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 826 -----AGGTACACAAATAATTAGAAAGAA 849  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 850 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTCTATAGTGCTGCTTTC 909  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 910 CTGCTGATCTATGATCTTATGCTGTGGCTTCTGTGATGGGACCACTTGGTCTCTCA 969  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 970 GGGGAA-----TATTCTATTGGCAAGTACTCCTGTTATTTCTTTCTGTTATTAAATGGG 1023  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaAtgGlyAla 320  
Db 1024 GCTTTTAGTGTGGACAGGATCTCCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGCA 1083  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1084 GCTTATCAAAATCTTCAAGATAATTGATAAAGCCAAAGTATTGACAGCTATTTCGAAGAGT 1143  
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1144 GGCACAAACCATGATATATTAAAGGGAATTTTGAATTTGCAAAATGTTCACTTCAGTTAC 1203  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLysAsnLeuArgIleLysSerGlyGlu 380  
Db 1204 CCATCTCGAAAGCAAGTAAAGATCTTGAAGGGCTTGAACCTGAAGGTGCAGAGTGGGCG 1263  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1264 ACAGTGGCTGTTGGAAACATGCTGTGGGAAGAGCACACAGTCCAGCTGATGCAG 1323  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1324 AGCTCTATGACCCACAGAGGGGATGGTGTGTTGATGGACAGGATATTAGACCATA 1383  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1384 AATGTAAGGTTTCTACGGGAATCATTTGGTGTGTCAGTCAGGAACCTGTTGTTGTC 1443  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1444 ACCACGATGCTGAAACATTCGCTATGGCCGTGAAAAATGTCACCATGGATGAGATTGAG 1503  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1504 AAAGCTGTCAGGAAGCAATAGCTTATGACTTTATCATGAATGCTCCTATAAATTTGAC 1563  
Qy 481 ThrLeuValGlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1564 ACCTGGTGGAGAGAGAGGGGCCAGTGTGAGTGTGGGAGAGAGAGAGATCGCCATT 1623  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1624 GCAGTGCCTGTTGGCAACCCCAAGATCCTCTGCTGATGAGGCCACGTCAGCCCTTG 1683  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
Db 1684 GACACAGAAAGCAAGCAGTGGTTTCAGTGGCTCTGATTAAG 1725

## RESULT 15

E02326  
 LOCUS 4378 bp RNA linear PAT 29-SEP-1997  
 DEFINITION Multidrug resistance relating gene derived from human normal cells.  
 ACCESSION E02326  
 VERSION E02326.1 GI:2170561  
 KEYWORDS JP 1990100860-A/1.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 4378)  
 Ueda.K. and Komano.T.  
 HUMAN NORMAL CELL-DERIVED MDR RELATED GENE  
 Patent: JP 1990100680-A 1 12-APR-1990;  
 SUNTORY LTD

## COMMENT

OS Homo sapiens  
 PN JP 1990100680-A/1  
 PD 12-APR-1990  
 PF 05-OCT-1998 JP 1989251475  
 PI UEDA KAZUMITSU, KOMANO TORU  
 PC C12N15/12, C12N1/21, C12Q1/68;  
 CC strandedness: Single;  
 CC topology: Linear;  
 CC \*source: tissue type=Adrenal gland;  
 CC \*source: clone=SAM132;  
 FH Key Location/Qualifiers  
 FH 5'UTR 1..137  
 FT CDS 138..3980  
 FT /gene='Multidrug resistance relating gene' FT  
 mat\_peptide 138..3977  
 FT /gene='MDR1',  
 FT 3'UTR 3981..>4378.

## FEATURES

source  
 1..4378  
 /organism='Homo sapiens'  
 /db\_xref='taxon:9606'  
 BASE COUNT 1315 a 818 c 1057 g 1188 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,33e-117 Length: 4378  
 Score: 1555.50 Matches: 306  
 Percent Similarity: 73.29% Conservative: 100  
 Best Local Similarity: 55.23% Mismatches: 99  
 Query Match: 56.79% Indels: 49  
 DB: 6 Gaps: 6

US-09-873-409-7 (1-541) x E02326 (1-4378)

QY 1 MetileuGlyleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 |||:|||||  
 DB 288 ATGGTGGTGGGAATTTGGCTGGCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 347  
 |||:|||||  
 QY 21 ValileuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 |||:|||||  
 DB 348 GTGTTGGAGAAATGACAGATATCTTTGGAATGCGAGAAATTTAGAAATCTGATGCTCA 407  
 |||:|||||  
 QY 32 GlyCysLeuValGlnThr-----AsnThrTy-SerPhePhe----- 43  
 |||:|||||  
 DB 408 AACATCACTAATAGAGTATATCAATGATACAGGGTTCCTTCATGAATCTGGAGGAGAC 467  
 |||:|||||  
 QY 44 -----ArgLeuThrLeuTyTyValGlyleGlyValAlaAlaLeuIlePheGlyTy 61  
 |||:|||||  
 DB 468 ATGACAGGATGCTATTATTACAGTGGAAATGGTCTGGGGTGGTGGTGGCTTAC 527  
 |||:|||||  
 QY 62 IleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 |||:|||||  
 DB 528 ATTCAGGTTTCATTTGGTGGCTGGAGCTGGAGACAAATACACAAATTTAGAAACAG 587  
 |||:|||||  
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
 |||:|||||

DB 588 TTTTTCATGCTATATATCGACAGGATAGCTGTTGATGTGCACCATGTGTGGGAG 647  
 |||:|||||  
 QY 102 LeuAsnThrArgMetThr--AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 |||:|||||  
 DB 648 CTTAACACCCGACTTACAGATGCTCCAGATTAATGAAGGAATTTGGTGAACAAAT 707  
 |||:|||||  
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 |||:|||||  
 DB 708 GGAATGTTCTTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 767  
 |||:|||||  
 QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 |||:|||||  
 DB 768 GGTGGAGCTAACCTTGTGATTTTGGCCATCAGTCTGTTCTTGGACTGTGAGCTGCT 827  
 |||:|||||  
 QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180  
 |||:|||||  
 DB 828 GTCTGGGCAAGATACATCTTCACTTACTGATAAGAACTCTTAGCGTATGCAGAAAGCT 887  
 |||:|||||  
 QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 |||:|||||  
 DB 888 GGAGCAGTAGCTAGCTGACAGGCTCTGGCAGCAATTAGAACTGTGATTGCAATTGGAGGACAA 947  
 |||:|||||  
 QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyTrpPheTyPhe 220  
 |||:|||||  
 DB 948 AAGAAAGAACTTGAA----- 962  
 |||:|||||  
 QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyThrGlnAsnLeuLysAsp 240  
 |||:|||||  
 DB 963 -----AGGTACACCAAAATTTAGAGAA 986  
 |||:|||||  
 QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260  
 |||:|||||  
 DB 987 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTGCTTTC 1046  
 |||:|||||  
 QY 261 PhePheMetAsnGlyTyThrTyGlyLeuAlaPheTyTyGlyThrSerLeuIleLeuAsn 280  
 |||:|||||  
 DB 1047 CTGCTGATCATGCTATCTTATGCTCTGGCTTCTGTTATGGGACCACTTGGTCTCTCA 1106  
 |||:|||||  
 QY 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 |||:|||||  
 DB 1107 GGGGAA-----TATCTATTGGCAAGTACTCCTGTTTCTGTTATTAATGGG 1160  
 |||:|||||  
 QY 301 SerTyCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
 |||:|||||  
 DB 1161 GCTTTTATGTTGGACAGCATCTCCAGCATTTGAAGCATTTGCAAAATGCAAGAGGACCA 1220  
 |||:|||||  
 QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
 |||:|||||  
 DB 1221 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAATTTGACAGCTATTGCAAGAGT 1280  
 |||:|||||  
 QY 341 GlyTyLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTy 360  
 |||:|||||  
 DB 1281 GGCAACAAACAGATATATTAAGGAAATTTGGAATTCAGAAATGTCATCTTCAGTTAC 1340  
 |||:|||||  
 QY 361 ProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
 |||:|||||  
 DB 1341 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGCTCTGAACTGGAAGTGCAGAGTGGGAG 1400  
 |||:|||||  
 QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
 |||:|||||  
 DB 1401 ACAGTGGCTGTTGGAAACAGTGGCTGTTGGAGAGACACACAGTCCAGCTGATGTCAG 1460  
 |||:|||||  
 QY 401 ArgLeuTyAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
 |||:|||||  
 DB 1461 AGGCTCTATGACCCACAGAGGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1520  
 |||:|||||  
 QY 421 AsnValArgHisTyArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
 |||:|||||  
 DB 1521 AATGTAAGTTTCTACGGGAATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1580  
 |||:|||||  
 QY 441 ThrThrIleSerAsnAsnIleLysTyGlyArgAspValThrAspGluGluMetGlu 460  
 |||:|||||

Db 1581 ACCACGATAGCTGAAACATTTCGCTATGGCCGTGAAATGTCACCATGGATGAGATTGAG 1640  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1641 AAGCTGTCAGGAGCCAAATGCCATATGACTTTATCATGAACACTGCCTCATAAATTGAC 1700  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1701 ACCCTGGTTGGAGAGAGAGGGGCCCACTTGAGTGGTGGGCAGAGCAGAGGATCGCCATT 1760  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1761 GCACGTGCCCTGGTTCGTAACCCCAAGATCCTCCTGCTGATGAGGCCACGTCAGCCTTG 1820  
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1821 GACACAGAAAGCGAAGCAGTGGTTTCAGGTGGCTCTGGATAAG 1862

Search completed: March 31, 2003, 02:23:23  
Job time : 7369.73 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 528.737 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739

Sequence: 1 MILGILASLVNGACPLMPL.....SESKSAVOALEKDTPTYSF 541

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2195239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame\_p2n.model -DEV=xlh  
-Q/cn2\_1/USFT0.spool/US09873409/runat\_27032003.115416.19216/app\_query.fasta\_1.7544  
-DB=N Geneseq\_101002 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN.1.1.4247 @runat\_27032003.115416.19216 -NCPU=6 -ICPU=3  
-NO\_XUPXY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=10 -XGAPEXT=0.5 -FGAPOPEXT=6 -FGAPEXT=7  
-YGAPOPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_101002.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2437	89.0	3699	24	ABK83223	Human transporter
2	1566.5	57.2	4189	21	AAZ49334	Murine multidrug r
3	1566.5	57.2	4189	24	ABA94367	Mouse BCRP DNA rel
4	1566.5	57.2	4313	14	AAQ38950	Mouse multidrug re
5	1561.5	57.0	4279	22	AAQ38950	Dog P-glycoprotein
6	1557.5	56.9	4279	22	AAQ38950	Dog P-glycoprotein
7	1557.5	56.9	4279	22	AAQ38950	Dog P-glycoprotein
8	1557	56.8	4195	22	AAF86128	Cynomologous monke
9	1555.5	56.8	3860	21	AAZ49332	Human wild-type mu
10	1555.5	56.8	3860	24	ABA94365	Human BCRP DNA rel
11	1552.5	56.7	4233	21	AAZ90198	Rat mdrlb2 (multis
12	1552.5	56.7	4233	22	AAF27498	Rat mdrlb2 multidr
13	1552.5	56.7	4279	22	AAQ34888	Dog P-glycoprotein
14	1552.5	56.7	4317	22	AAQ34889	Dog P-glycoprotein
15	1552.5	56.7	4669	19	AAV32645	Human P glycoprote
16	1552.5	56.7	4669	24	ABK52041	cDNA encoding huma
17	1548.5	56.5	4186	22	AAF86127	Cynomologous monke
18	1548	56.5	4254	24	ABK63517	Rat sequence diffe
19	1546.5	56.5	3860	21	AAZ49333	Human G185V mutant
20	1546.5	56.5	3860	24	ABA94366	Human BCRP DNA rel
21	1546.5	56.5	4349	22	AAH57442	Human intestine ce
22	1546.5	56.5	4646	15	AAQ72872	Human multidrug re
23	1546.5	56.5	4646	21	AAZ94738	Human ATP binding
24	1546.5	56.5	4646	24	AAZ38994	Human mdrl gene.
25	1546.5	56.5	4646	24	ABL68592	Kidney cancer rela
26	1546.5	56.5	4646	24	ABL68880	Kidney cancer rela
27	1546.5	56.5	6505	17	AAZ13394	Hybrid vector p8F-
28	1546.5	56.5	8630	21	AAZ24041	Retroviral M4 mdr-
29	1546.5	56.5	8630	21	AAZ24042	Retroviral vector
30	1545	56.4	4264	19	AAV65533	Mutated human P-gl
31	1545	56.4	4264	19	AAV65534	Mutated human P-gl
32	1544.5	56.4	4669	8	AAZ04522	Sequence of human
33	1543.5	56.4	4378	11	AAQ04522	Multidrug Resistan
34	1543	56.3	4425	21	AAZ52048	Rat multidrug resi
35	1541	56.3	4369	21	AAZ52047	Rat multidrug resi
36	1540.5	56.2	4669	14	AAQ52726	Sequence of human
37	1539	56.2	3840	24	ABL91687	Human polynucleoti
38	1539	56.2	3988	21	AAZ88973	Human MDR-1 DNA.
39	1536	56.1	4788	21	AAZ49335	Murine multidrug r
40	1536	56.1	4788	24	ABA94368	Mouse BCRP DNA rel
41	1455.5	53.1	3924	21	AAZ94742	Human ATP binding
42	1455.5	53.1	3924	21	AAZ88974	Human MDR-3 DNA.
43	1455.5	53.1	3924	24	ABN95801	Gene #2299 used to
44	1421.5	51.9	3912	24	ABK63653	Rat sequence diffe
45	1350	45.6	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1	
ABK83223	
ID	ABK83223 standard; cDNA; 3699 BP.
XX	ABK83223;
AC	ABK83223;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX	
XX	Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW	neurological disorder; muscle disorder; immunological disorder; cancer;
KW	scleleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW	cell proliferative disorder; cervical cancer; breast cancer;
KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW	myotonic dystrophy; catatonias; endocrine disorder; diabetes;
KW	Grave's disease; gastrointestinal disorder; Crohn's disease;
KW	renal disorder; Good pasture's syndrome; viral infection; cirrhosis;



KW bacterial infection; fungal infection; parasitic infection;  
 KW protozoal infection; helminthic infection; cardiovascular disorder;  
 KW atherosclerosis; hepatic disease.

OS Homo sapiens.

XX WO200240541-A2.

XX 23-MAY-2002.

XX 25-OCT-2001; 2001WO-US46055.

XX 27-OCT-2000; 2000US-243989P.

XX 03-NOV-2000; 2000US-245904P.

XX 09-NOV-2000; 2000US-247673P.

XX 17-NOV-2000; 2000US-249661P.

XX 20-DEC-2000; 2000US-252232P.

XX 01-DEC-2000; 2000US-250790P.

XX (INCYTE GENOMICS INC.

XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;

XX Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;

XX Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimza Y, Khan FA;

XX Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;

XX Ison CH, Das D, Raumann BE, Policky JL, Kearney L;

XX WPI; 2002-463570/49.

XX P-PSDB; ABG61544.

XX New transporters and ion channels (TRICH) polypeptides, useful for  
 PT diagnosing, preventing, and treating disorders associated with an  
 PT abnormal expression or activity of TRICH, e.g. immunological, muscular  
 PT or renal disorders

XX Claim 5; Page 172; 178pp; English.

XX The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
 CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
 CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
 CC a recombinant polynucleotide comprising a promoter sequence operably  
 CC linked to the TRICH polynucleotide, a cell transformed with the  
 CC recombinant polynucleotide, a transgenic organism comprising the  
 CC recombinant polynucleotide, an isolated antibody that binds specifically  
 CC to TRICH, and screening for compounds which bind to TRICH, modulate  
 CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.

CC The polypeptides are useful for diagnosing, treating, and  
 CC preventing transport, neurological, muscle, immunological disorders  
 CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
 CC proliferative disorders such as cancers (e.g. leukaemia, cervical or  
 CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
 CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
 CC catatonla), endocrine disorders (e.g. diabetes, Grave's disease),  
 CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
 CC (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,  
 CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
 CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
 CC other diseases and disorders detailed in the specification. They can also  
 CC be used in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of transporters and  
 CC ion channels. TRICH or its fragments may also be used in screening for  
 CC compounds that specifically bind to and modulate the activity of TRICH.  
 CC The polynucleotides can be used to create knock-in humanised animals or  
 CC transgenic animals to model human disease. The present sequence  
 CC encodes a TRICH protein.

XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:

Fred. No.: 1.45e-241 Length: 3699  
 Score: 2437.00 Matches: 495  
 Percent Similarity: 90.71% Conservative: 3

Best Local Similarity: 90.16% Mismatches: 9  
 Query Match: 88.97% Indels: 42  
 DB: 24 Gaps: 3

US-09-873-409-7 (1-541) x ABK83223 (1-3699)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 145 ATGATCCTGGGTATACCTGACATCAGTCTTCAATGGAGCCTCCCTTCTTAATGCCACTG 204  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
 DB 205 TGTATAGGAGAAATAGTATACCTTATTAGTGGATGCTAGTCCACACACAACA 264  
 QY 41 SerPhePheArg-----LeuThr 46  
 DB 265 AATTATCAGAACTGTACTCAGTCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324  
 QY 47 LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
 DB 325 CTGTATTATGTTGGATAGTGTGCTGCTTGTATGTTTGGTTACATACAGATTTCCTTG 384  
 QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86  
 DB 385 TGGATTATTAACCTGCAGCAGCAGACAAAGAGATTTCGAAACAGATTTTTCATTTCAGTT 444  
 QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyLeuLeuAsnThrArgMet 106  
 DB 445 TTGGCAGACAGACATCGCTGGTGTGATGCTGTGATCGTGAACCTTAACATCGCATG 504  
 QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
 DB 505 ACAGATGACATTGACAAATCAGTGTGTTGGAGATAAGATTCTCTGTTGTTTCAA 564  
 QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145  
 DB 565 AACATGCTACTATTTTCGATTGGCTGGCAGTTGGTGGTGAAGGGCTGGAAACCTCACC 624  
 QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165  
 DB 625 CTAGTGACTCTATCCAGCTCTCTCTTATTAATGGCTTCAGCGCAGCATGTTCTAGATG 684  
 QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
 DB 685 GTCATCTCTATGACCACTAAGGAATTAAGTGCCTATTCCAAGCTGGGCTGGGAGAA 744  
 QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205  
 DB 745 GAAGTCTTGTCTCAATCCGACAGTCATAGCCTTTAGGGCCCGAGAGAAAGAACTTCAA 804  
 QY 206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTTrpLeuLeu 225  
 DB 804 ----- 804  
 QY 226 SerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245  
 DB 805 -----AGGTATACACAGAACTCAAGATCAAGAGATTTCG 843  
 QY 246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265  
 DB 844 ATAAAAAGGACTATAGCTCAAAAGTGTCTCTTGGTGGCTGTGTACTTCTTTATGAATGGA 903  
 QY 266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285  
 DB 904 ACCTATGGACTTCTCTTTTGGTATGGAACCTCTCTGATTTCTTATGAGAACCTGGATAT 963  
 QY 286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305  
 DB 964 ACCATCGGAGCTGTTCTTCTGCTGTTTCTTTAGTGTAAATCCATAGTAGTTATTGCATTGGA 1023  
 QY 306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325  
 DB 1024 GCAGCAGTCCCTCCTTGAACCTTTCGAATAGCCCGAGGAGCTGCTCTTCATATTTTC 1083

QY 326 GlnValIleAspLysLysProSerIleAspAsnPhSerThrAlaGlyTyrLysProGlu 345  
 DB 1084 CAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCACAGCTGGATATAACCTGAA 1143  
 QY 346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365  
 DB 1144 TCATAGAAGGAAGTGTGAATTTAAATATGTTCTTTCAATTTATCCATCAAGGCCATCT 1203  
 QY 366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385  
 DB 1204 ATCAAGATTCTGAAGGTCTGATCTCGGAATTAAGTCTGGAGACAGCTCGCCTTGGTC 1263  
 QY 386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLysTyrAspPro 405  
 DB 1264 GGTCTCAATGGCAGTGGGAAGAGTAGCTGAGCTTCTGCAGAGGTTATATGATCCG 1323  
 QY 406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425  
 DB 1324 GATGATGGCTTTATCATGTTGGTGAATGAGATGACATCAGAGCTTTAAATGTGGCGCATTTAT 1383  
 QY 426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445  
 DB 1384 CGAGACCATATTGGATGGTGTAGTCAAGAGCCTGTTTGTTCGGGACCACCATCAGTAAC 1443  
 QY 446 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465  
 DB 1444 AATATCAAGTATGCGAGAGATGATGTGACTGTAGTGAAGAGATGAGAGAGCAGCAAGGAA 1503  
 QY 466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485  
 DB 1504 GCAAAATCGCTATGATTTTATCATGGAGTTCCTTAATAAATTTAATACATTTGGTAGGGAA 1563  
 QY 486 GlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505  
 DB 1564 AAGGAGCTCAAAATGATGGAGGCGAGAAACAGAGATCGCAATTCCTGTCCTTAGTT 1623  
 QY 506 ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525  
 DB 1624 CGAAACCCCAAGATTCGATTTTAGATGAGGTACGTCGCTCGGATTCAGAAAGCAAG 1683  
 QY 526 SerAlaValGlnAlaLeuGluLys 534  
 DB 1684 TCAGCTGTTTCAGCTGCACTGGAGAAG 1710  
 RESULT 2  
 AAZ49334  
 ID AAZ49334 standard; cDNA; 4189 BP.  
 AC AAZ49334;  
 XX  
 XX 14-MAR-2000 (first entry)  
 DE Murine multidrug resistance-1 (MDR-1) cDNA.  
 KW Multidrug resistance; MDR-1; P-glycoprotein;  
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassaemia;  
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
 KW cytokine; ds.  
 XX  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..3831  
 CDS /tag= a  
 FT /product= "Murine MDR-1 protein"  
 FT  
 FT  
 FT  
 XX  
 PN W09961589-A2.  
 XX  
 XX  
 PD 02-DEC-1999.  
 XX

PF 27-MAY-1999; 99WO-US11825.  
 XX  
 PR 28-MAY-1998; 98US-0086988.  
 XX  
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Sorrentino B, Bunting K;  
 XX  
 DR WPI; 2000-072615/06.  
 DR P-PSDB; AAY58188.  
 PT Ex vivo expansion of hematopoietic stem cells transduced with a  
 PT sequence encoding human multidrug resistance-1, used for bone marrow  
 PT transplantation -  
 XX  
 PS Disclosure; Page 90-93; 113pp; English.  
 XX  
 CC This sequence represents cDNA encoding murine multidrug  
 CC resistance protein MDR-1. MDR-1 is a transmembrane  
 CC efflux pump, responsible for the export of drugs from cells,  
 CC particularly cancer cells. The invention relates to transducing  
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified haematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC haematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,  
 CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in haematopoietic stem cells.  
 CC Haematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC haematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.  
 XX  
 SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,06e-151 Length: 4189  
 Score: 1566.50 Matches: 307  
 Percent Similarity: 72.63% Conservative: 99  
 Best Local Similarity: 54.92% Mismatches: 94  
 Query Match: 57.13% Indels: 59  
 DB: 21 Gaps: 6  
 US-09-873-409-7 (1-541) x AAZ49334 (1-4189)  
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20  
 DB 148 ATGATTCGGGAACCTCGCTGCTATTATCCATGGNACATTACTCCCTCTTGATGCTG 207  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
 DB 208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACGAGTATTCTGCAAGCATT 267  
 QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
 DB 268 ACTAATCAAAGTGGACCCCAACAGTAGTCTCTGATCATCAGCAACAGCAGTCTGGAGGAGAG 327  
 QY 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValaAla 56  
 DB 328 ATGGCCATATAGCCCTAC-----TATTACCGGGATTTGCTGCTGTG 372  
 QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76  
 DB 373 CTCATAGTTGCTTACATCCAGGTTTCACCTTGTGCTGCGAGCTGGAAGACAGATACAC 432  
 QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
 DB 433 AAGATTAGGCAGAAAGTTTTTTCATGCTATAATCAATCAGGAGATAGGCTGGTTGATGTG 492

QY 97 CysAspIleGlyGluLeuAenThrAtqMetThr---AspIleAspLysIleSerAspGly 115  
Db 493 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGTCCTCCAAATAATATGACGGA 552  
QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
Db 553 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCAATTTTTAGCCGGTTTTATC 612  
QY 136 ValGlyLeuValLysGlyTyrLysLeuThrIleuValThrLeuSerThrSerProLeuIle 155  
Db 613 ATAGGATTTATAAGTGGTGGAGCTAACCTTGTCTATTTTGGCTGTACGCCCTCTTATT 672  
QY 156 MetAlaSerAlaAlaAlaCysSerAtqMetValIleSerLeuThrSerLysGluLeuSer 175  
Db 673 GGATTGTCATCTGTTTGGGCAAGAGTATTGATCTTACTATTACTATAAGGAATCCAG 732  
QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195  
Db 733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792  
QY 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuValThrArgTyr 215  
Db 793 GCCTTTGGAGGACAAAGAGAACTTGAA-----822  
QY 216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu---PheValArgTyrThr 235  
Db 823 -----AGTACAAAT 831  
QY 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255  
Db 832 AAAAATTTAGAGAAGCTAAATAATGTTGGCATAAAGAAGCTATCACAGCCAGCAATTCG 891  
QY 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThr 275  
Db 892 ATAGGCATTGCTACCTGTTGGTCTATGTCATCATATGCTGTCATTCCTGTTATGGGACA 951  
QY 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295  
Db 952 TCCTTGGTCTCTCAAAATGAA-----TATTCTATTGAGAAGTCTTACTGCTCTCTC 1005  
QY 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315  
Db 1006 TCTATTGTTGGGACTTTTAGTATTGGACACTTGGCCCCAACATAGAAAGCTTTGCA 1065  
QY 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335  
Db 1066 AACGCACGAGGGGCGAGCTTTGAAATCTTCAAGATAATTGATAACGAGCCCAAGCAATTGAC 1125  
QY 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355  
Db 1126 AGCTTCTCAAAAGGGCTCAAAACAGACAGATATAATGGAAACCTTAGAGTTTAAAAAT 1185  
QY 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375  
Db 1186 GTTCACCTTCAACTACCATCGAAGGGAAGTTCAGATCTTGAAGGCGCTCAATCTGAAG 1245  
QY 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395  
Db 1246 GTGAAGAGCGGACAGACGGTGGCTTGGTGGCAACAGTGGCTGTGGAAAAAGCAACAAT 1305  
QY 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415  
Db 1306 GTCCAGCTGATGACAGAGCTCTACGACCCCTGGAGGGCGTGGTCAATTCGACGGACAA 1365  
QY 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435  
Db 1366 GACATCAGAACCATCATGTGAGGTATCTGAGGAGATCATTTGGTGGTGTGAGTCAGGAA 1425  
QY 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr 455  
Db 1426 CCTGTGCTGTTTGGCCACGATCGCCGAGAACCATTCGCTATGCGCGAGAGATGTCACC 1485

QY 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475  
Db 1486 ATGGATGAGATTGAGAAAGCTGCTCAAGGAAGCAATGCTATGACTTTCATCATGAAACTG 1545  
QY 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495  
Db 1546 CCCACCAATTTGACACCCTGTTGTTGAGAGAGGGCGCAGTGAAGTGGGGGACAGAAA 1605  
QY 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515  
Db 1606 CAGAGAATCGCATTCGCCGGCCCTGGTGGCCCAATGCCCAAGATCCTTTTGTGGACGAG 1665  
QY 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1666 GCCACCTCAGCCTGGATACAGAAAGTGAAGCTGTGTCAGGCGCCACTGGATAAG 1722  
RESULT 3  
ABA94367  
ID ABA94367 standard; DNA; 4189 BP.  
XX ABA94367;  
AC ABA94367;  
XX 26-MAR-2002 (first entry)  
XX Mouse BCRP DNA related seq Id No. 5.  
DE Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;  
KW hepatotropic; neurodegenerative; cytosstatic; antianemic; muscular; BCRP;  
KW cardiant; gene therapy; ds.  
XX Mus musculus.  
FH Key Location/Qualifiers  
FT CDS 1..3831  
FT /\*tag= a  
XX WO200192877-A2.  
XX 06-DEC-2001.  
XX 30-MAY-2001; 2001WO-US17459.  
XX 31-MAY-2000; 2000US-0584586.  
PR 29-MAY-2001; 2001US-0866866.  
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
PA Sorrentino B, Schuetz J;  
XX WPI; 2002-114368/15.  
DR P-PSDB; ABB07268.  
XX Identifying a stem cell, for treating e.g., muscular dystrophy,  
PT myocardial infarction, Parkinson's disease, or neurodegenerative  
PT disorders, comprises detecting the expression of an ATP transport  
PT protein (BCRP) by a cell  
XX Disclosure; Page 64-66; 87pp; English.  
XX The invention provides a method of identifying and/or isolating a stem  
CC cell that involves detecting the expression of an ATP transport protein  
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample  
CC comprising stem cells. The isolated stem cells may be used in the  
CC treatment of diseases such as muscular dystrophy, degenerative liver  
CC disorder, myocardial infarction, Parkinson's disease, degenerative  
CC disorders of the brain, and for tissue regeneration or replacement.  
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for  
CC treatment of leukemia) and for ex vivo gene therapy for treating blood  
CC diseases such as sickle cell anemia and thalassemia. The stem cells can  
CC also be used as cell targets in gene therapy protocols. The present  
CC sequence represents a sequence related to the BCRP for which no relevant  
CC information has been provided in the specification.  
XX

SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

## Alignment Scores:

Pred. No.: 2,06e-151 Length: 4189  
Score: 1566.50 Matches: 307  
Percent Similarity: 72.63% Conservative: 99  
Best Local Similarity: 54.92% Mismatches: 94  
Query Match: 57.19% Indels: 59  
DB: 24 Gaps: 6

US-09-873-409-7 (1-541) x ABA94367 (1-4189)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 148 ATGATTCTGGGAACCTCTCGTGTCTATTATCATGGAACATTACTTCCCTCTTTGATGCTG 207  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
DB 208 GTGTTTGGAAACATGACATAGTATTACAAAGCAGAACCCAGTATTCTGCCAAGCATT 267  
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
DB 268 ACTAATCAAGTGGACCCACACACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327  
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56  
DB 328 ATGCCATATACGCCTAC-----TATTACACCGGGATTGGTGTGTG 372  
QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76  
DB 373 CTATAGTTCCTTACATCCAGGTTTCACTTGTGTGCTGGCAGCTGGAACAGATACAC 432  
QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
DB 433 AAGATTAGGCAGAGATTTTTCATGTCTATATGATCAATCAGAGATAGCTGGTTGATGTC 492  
QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspIleSerAspGly 115  
DB 493 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGTCCTCCAAAATTAATGACGGA 552  
QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
DB 553 ATTGGTGACAAAATTGGGATGTTTTCAGTCCATACCAACCATTTTAGCCGTTTATC 612  
QY 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
DB 613 ATAGATTATTAAGTGTGGAGCTNACCCTTGTCTATTTGGCTGCACCCCTCTATT 672  
QY 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
DB 673 GGATTGTCTATGCTTGTGGCAAGGTATTGACTTTCATTACTAATAAGGAACCTCCAG 732  
QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195  
DB 733 GCTTATGCAAAAGCTGGAGCAGTGTCTGGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792  
QY 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215  
DB 793 GCCTTTGGAGGACACAGAAGAACTTGAA----- 822  
QY 216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThr 235  
DB 823 -----AGGTACAAT 831  
QY 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255  
DB 832 AAAAAATTAGAAGAGCTAAATAATGTTGGCAATAAGAAAGCTATATCAGCCAGCATTCG 891  
QY 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275  
DB 892 ATAGGCAATGCCTACCTGGTGGTATGCAATCATATGCACTGGCATTTCTGGTATGGACA 951  
QY 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295

DB 952 TCTTGGTCTCTCAAATGAA-----TATTCTATTGGAGAAGTGTCTACTGTCTCTTC 1005  
QY 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315  
DB 1006 TCTATTTTTGGGACACTTTTAGTATTGGACACTTGGCCCCCAACATAGAGCCTTTGCA 1065  
QY 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335  
DB 1066 AAGCCACGAGGGGAGCCCTTTGAAATCTTCAAGATAAATGATAACGAGCCAGCATTGAC 1125  
QY 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355  
DB 1126 AGCTTCTCAACAAAGGGCTACAAACACAGACGTATAATCGGAACCTTAGAGTTTAAAAAT 1185  
QY 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375  
DB 1186 GTTCACTTCACTACCCATCGAGAGCGAAGTTCAGATCTTGAAGGCGCTCAATCTGAAG 1245  
QY 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395  
DB 1246 GTGAAGAGCGGACAGACGGTGGCTTGGTTGGCAACAGATGGCTGTGGAACAAACAACT 1305  
QY 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415  
DB 1306 GTCCAGCTGATCGAGGCTCTACGACCCCTGGAGGGCGGTGTCAGTATCGACCGACAA 1365  
QY 416 AspIleArgAlaAlaAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435  
DB 1366 GACATCAGAACCATCATNTGTAGGTATCTGAGGGAGATCATTTGGTGTGTGATCAGGAA 1425  
QY 436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThr 455  
DB 1426 CCTGTGCTGTTGCCACCATCGATCGCCGAGAACATTCGCTATATGGCCGAGAGATGTCA 1485  
QY 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475  
DB 1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCCAATGCCCTATGACTTCATCATGAA 1545  
QY 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495  
DB 1546 CCCCACCAATTTGACACCTTGGTTGGTGGAGAGGGCCAGCTGAGTGGGGACAGAAA 1605  
QY 496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515  
DB 1606 CAGAGAATCGCCATTCGCCGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTGGACGAG 1665  
QY 516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
DB 1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGTGTCAGGCCGCACTGGATAAG 1722

## RESULT 4

AAQ38950  
ID AAQ38950 standard; DNA; 4313 BP.

XX AAQ38950;

DT 28-JUL-1993 (first entry)

DE Mouse multidrug resistance sequence.

KW mdr gene; Lambda DR11 clone; ss.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 110..3940

FT /\*tag= a

FT /phenotype= multidrug\_resistance

XX US5198344-A.

XX 30-MAR-1993.

PD

```
XX 15-JUL-1986; 86US-0885951.
XX
XX 15-JUL-1986; 86US-0885951.
XX 06-FEB-1991; 91US-0852311.
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Croop JM, Gros P, Housman DE;
XX
XX WPI; 1993-126077/15.
XX P-PSDB; AAR35199.
XX
XX DNA sequence which confers multi-drug resistance on sensitive
XX mammalian cells - used to preserve bone marrow cells during
XX chemotherapy to prevent infection
XX
XX Claim 1; Fig 8; 22pp; English.
XX
XX A cDNA library was constructed from mRNA which had been isolated
XX from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
XX complementary to the mRNA species encoded by 2 related but distinct
XX mdr genes were isolated and cloned. One of the cDNA clones (lambda
XX dril) is a full-length cDNA clone for one member of the mdr gene
XX family. When incorporated into prokaryotic expression vector pDREX4
XX (which allows high levels of transcription of the cDNA when
XX introduced into mammalian cells), the clone was shown to confer the
XX multidrug resistance phenotype upon transfection into drug-sensitive
XX mammalian cells.
XX
XX SQ Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:
Pred. No.: 2.15e-151 Length: 4313
Score: 1566.50 Matches: 307
Percent Similarity: 72.63% Conservative: 99
Best Local Similarity: 54.92% Mismatches: 94
Query Match: 57.19% Indels: 59
DB: 14 Gaps: 6

US-09-873-409-7 (1-541) x AAQ38950 (1-4313)
Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 257 ATGATTCTGGGAACCTCGCTGCTATTATCCATGGAACATTACTTCCCTCTTGATCGT 316
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30
Db 317 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAACCCAGTATTCTGCCAAGCATT 376
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 377 ACTAATCAAAAGTGACCCCAACAGACTCTGATCATCATGCAACAGCAGCTCTGGAGGAAGAG 436
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 437 ATGCCCATATAGCCTAC-----TATTACACCGGGATTTGGTGGTGTG 481
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLys 76
Db 482 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGGCTGGCAGCTGGAAGACATACAC 541
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db 542 AAGATTAGGCGAAGATTTTCCATGCTATATGAATCAGGAGATAGGCTGGTTGATGTG 601
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db 602 CATGATGTTGGGAGCTCAACACCGCGCTCACAGATGATGCTCCAAATTAATGACGGA 661
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db 662 ATTGCTGACAAAATTTGGGATGTTTTTTTTCAGTCCATACCAACCATTTTATGCGCGT 721
```

```
Qy 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db 722 ATAGGATTATTAAGTGGTTGGAAAGCTAACCCTTGTTCATTTTGGCTGTGAGCCCTCTTATT 781
Qy 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db 782 GGATTGTCATCTGCTTTGTGGCAAAAGGTAATGACTTCATTACTAATAAGGAATCCAG 841
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195
Db 842 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAAGAACTGGATT 901
Qy 196 AlaPheArgAlaGlnGlnLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db 902 GCCTTTGGAGGACACAGAGGAACTTGAA----- 931
Qy 216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db 932 -----AGGTACAAT 940
Qy 236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db 941 AAAAATTTAGAAAGACTAAATAATGTGGCATAAAGAAAGCTATCACAGCCAGCATTTCCG 1000
Qy 256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThr 275
Db 1001 ATAGGCAATTCCTACCTGTTGTCATGTCATCATATGCACTGGCATTCGGTATGGGACA 1060
Qy 276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db 1061 TCCTTGGTCTCTCAAAATGAA-----TATTCTATTGGAGAAGTGTACTGCTCTCTTCTTC 1114
Qy 296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db 1115 TCTATTTTGTGGGACTTTTATGATTGGACACTTGGCCCCCAACACATAGAAGCCTTTGCA 1174
Qy 316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db 1175 AACGCACAGGGGCGACCTTTGNAATCTTCAAGATTAATTAATACAGCCAGCCAGCATTTGAC 1234
Qy 336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db 1235 AGCTTCTCAACAAAGGGCTACAAACACAGACAGTATTAATGGAAACTTAGAGTTTAAAAAT 1294
Qy 356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db 1295 GTTCACCTTCAACTTACCCTACCGAAGCGAAGTTTCAGATCTTGAAGGGCCCTCAATCTGAAG 1354
Qy 376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db 1355 GTGAAGAGCGGACAGACGGTGGCTTGGTTGGCAACAGTGGCTGTGTAAGAAACACCAACT 1414
Qy 396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db 1415 GTCCAGCTGATGCAGAGGCTCTACGACCCCTCGGAGGGGTGTGTATTCGACGAGACAA 1474
Qy 416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db 1475 GACATCAGAACCACTCAATCATGTAGGTATCTCGAGGGAAGTCAATGTGTGTGTGAGTCAAGAA 1534
Qy 436 ProValLeuPheGlyThrThrIleSerAsnIleLysTyrGlyArgAspValThr 455
Db 1535 CCTGTGCTGTTTGGCCACAGCATCGCCGAGACATTCGCTATGCGCCGAGAGATGTACACC 1594
Qy 456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db 1595 ATGATGATGATTGAGAAAGCTGTCAAGGAAGCAATGCTATGCTTTCATCATGAACATG 1654
Qy 476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlnLys 495
Db 1655 CCCACCAAAATTTGACACCTGGTTGTTGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1714
```



```
Db 926 GCTTCTTATTGATCTATGTCATCATATGCTCTGGCTTTCTGGTATGGACCTCCTCTGGTC 985
Qy 279 LeuAsnGlyGluProGlyThrValLeuAlaValPheSerValIle 298
Db 986 CTCTCCAGTGAA-----TATACATATTGGACAGTACTCACTGCTCTCTTTCTGTATTA 1039
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db 1040 ATTTGGGCTTTTAGTATTGGACAGCATCCCAAGCATTGAAGCATTGCAACGCAAGA 1099
Qy 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db 1100 GGAGCAGCTTATGAATCTTCAAGATATTGACAAATAAACAAGCATTGACAGCTATTTCG 1159
Qy 339 ThrAlaGlyThrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db 1160 AAGAGTGGACATAAACCATGATATTAAAGGAAATTTGGAATTCAAAAATGTTCACTTC 1219
Qy 359 AsnTyrProSerArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSer 378
Db 1220 AGTTACCTCTTCGAAAGAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTTCAGAGT 1279
Qy 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db 1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGTTCGCGGAAGAGCAGCAGCCGTGCAGCTG 1339
Qy 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db 1340 ATGCAGAGGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGACAGGACATTAGG 1399
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db 1400 ACCATAATGTAAAGCATCTTCGGGAATTAATCTGGTGTGGTGTGAGTCAGGAGCCTGTG 1459
Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db 1460 TTTGCCACACGATAGTGAACCAATTCGCTATGCGCGCAAAATGTCACCATGGATGAG 1519
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db 1520 ATTGAGAAAGCTGTAAAGGAAGCCCAATGCCCTATGATTTATCATGAACATACCTTAATAA 1579
Qy 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db 1580 TTTGACACTCTGGTTGGAGAGAGGGCCCGCAGCTGAGTGGTGACAGAAACAGAGATC 1639
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db 1640 GCCATTGCTCGGGCCCTGGTTGCGCAACCCCAAGATTCTTCTGCTGGATGAGCAACGTCA 1699
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
Db 1700 GCTCTGGACACTGAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

## RESULT 6

```
AAD03504
ID AAD03504 standard; cDNA; 4279 BP.
XX AAD03504;
XX AAD03504;
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
PH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
```

```
FT
FT (Genotype A) protein"
FT replace (607, C)
XX /*tag= b
PN WO200123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI; 2001-235373/24.
DR P-PSDB; AAE00308.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
XX Claim 9; Page 85-90; 11pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX
XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.
XX
XX Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;
```

## Alignment Scores:

```
Pred. No.: 1.81e-150 Length: 4279
Score: 1557.50 Matches: 303
Percent Similarity: 72.12% Conservative: 98
Best Local Similarity: 54.50% Mismatches: 104
Query Match: 56.86% Indels: 51
DB: 22 Gaps: 5
```

US-09-873-409-7 (1-541) x AAD03504 (1-4279)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTGGTGGGACAAATGGCTGCCATCATCCATGGAGCTGCATCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGTCAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePhearg----- 44
Db 287 CCAGTTATAATTAATGAAGTATTACGAACAATACACAACTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGAAATGACCACGATGCTGCTATTATTCAGTGGGATCGGTCTGGCTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
```



Db 407 GCTTACATCCAGGTTTCATTTCTGGTGCCTGGCAGCAGGAGACACATCTCAAAATTAGA 466  
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
Db 467 AAACAATTTTTTCATGCTATATCGCAGCAGGAGATTGGCTGGTTTGAAGCTGATGACGTT 526  
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAsp 118  
Db 527 GGGGAGCTTAACACCCGGCTCACACACATGCTCTCAAAATCAATGAAGGAATTGGCCAC 586  
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 587 AAATTTGAATGTTCTTTCATCATATGACACATTTTTCACCGGTTTATAGTGGGGTTT 646  
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 647 ACACGTGGTGGAGCTAAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACATTCA 706  
Qy 159 AlaAlaIaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 707 GCCGCCATCTGGGCAAGATATCTTTCATTTACTGATAAAGAACTCTTGGCCCTATGCA 766  
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 767 AAGCTGGAGCAGTAGCTGAAGAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 826  
Qy 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPhe 218  
Db 827 GGACAAAGAAAGAACTTGAA----- 847  
Qy 219 TyrPheProGlnTyrLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeu 238  
Db 848 -----AGTCAACAACAAAATTATA 865  
Qy 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
Db 866 GAAGAAGCTTAAGGAATTGGGATAAAGAAAGTCTACAGCCCAACAACTTCTATTGTTGCC 925  
Qy 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIle 278  
Db 926 GCTTTCATTATGATCATCATCATCTCTGGCTTCTGATGGGACCTCTTGGTC 985  
Qy 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
Db 986 CTCTCCAGTGAA-----TATTCTATTGGACAAGTACTCACTGTCTCTTCTTCTGATTATA 1039  
Qy 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318  
Db 1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCAAGCATTTGAAGCATTTTGAACACCAAGA 1099  
Qy 319 GlyAlaAlaPheHisPheIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338  
Db 1100 GGAGCAGCTTATGAATCTTCAAGATATTGCAATAAACCAAGCATTTGACAGCTATTG 1159  
Qy 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358  
Db 1160 AAGAGTGACATAAACCAAGATATAATTAAAGGAATTTGGAATTTCAAAATCTTCACTTC 1219  
Qy 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
Db 1220 AGTTACCTCTCTCGAAAGAAAGTAAAGATCTTAAAGGGTCTCAACCTGAAGGTTGAGAT 1279  
Qy 379 GlyGluThrValAlaLeuValGluLeuAsnGlySerGlyLysSerThrValValGlnLeu 398  
Db 1280 GGGCAGACAGTGGCGTGGTGGGAACAGTGGCTGGGGAAGAGCAGCACCGCTGCAGCTG 1339  
Qy 399 LeuGlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArg 418  
Db 1340 ATGCAGAGCTCTATGACCCACAGATGCGATGGTCTGTATTGATGAGCAGACATTAGG 1399  
Qy 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
Db 1400 ACCATAAATGTAAAGCATCTTCGGGAAATTAATTACTGGTGTGTTGAGTCAAGGAGCTGTGTTG 1459

Qy 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458  
Db 1460 TTTTGCCACCATAGCTGAAACATTCGCTATGCGCCGCAAAATGTCCCATGTGATGAG 1519  
Qy 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
Db 1520 ATTGAGAAGCTGTTAAGGAAGCCCATGCTATGATTTTATCATGAAACTACTTAATAA 1579  
Qy 479 PheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
Db 1580 TTTGACACTCTGTTGGAGAGAGAGGGGCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639  
Qy 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
Db 1640 GCCATTGCTGGGCCCTGTTTCGAACCCCAAGATTCTTCTGCTGATGAGGCAACGTC 1699  
Qy 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1700 GCTCTGGACACTGAAAGTGAAGCACTGTTTCCAGGTGGCCCTGGATNAG 1747  
RESULT 7  
AAD03505  
ID AAD03505 standard; cDNA; 4279 BP.  
XX  
AC AAD03505;  
DT 13-JUN-2001 (first entry)  
XX  
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.  
XX  
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
KW MDRI; drug bioavailability; transgenic animal; genetic model; ss.  
XX  
OS Canis familiaris.  
XX  
FH Key Location/Qualifiers  
FT CDS 17..3862  
FT /\*tag= a  
FT /product= "Dog P-glycoprotein (PGP) allelic variant  
FT (Genotype B) protein"  
FT replace (91, T)  
FT /\*tag= b  
FT allele replace (607, C)  
FT /\*tag= c  
XX  
PN WO200123540-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-US26767.  
XX  
PR 28-SEP-1999; 99US-0156510.  
XX  
PA (GENT-) GENTEST CORP.  
XX  
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
XX  
DR WPI; 2001-235373/24.  
DR P-PSDB; AAE00309.  
XX  
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT PGP inhibitors -  
XX  
PS Claim 9; Page 93-99; ilpp; English.  
XX  
CC The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDRI) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and

CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.

XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

## Alignment Scores:

Pred. No.: 1,81e-150 Length: 4279  
 Score: 1557.50 Matches: 303  
 Percent Similarity: 72.12% Conservative: 98  
 Best Local Similarity: 54.50% Mismatches: 104  
 Query Match: 56.86% Indels: 51  
 DB: 22 Gaps: 5

US-09-873-409-7 (1-541) x AAD03505 (1-4279)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 DB 167 ATGTTGGTGGGACATGCTGCATCATCCATGGAGTGCATCTCCTCATGATGCTG 226  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
 DB 227 GTTTTGGAAACATGACAGATAGCTTTGCAATGCGAGGAATTTCAAGAAACAAACTTTT 286  
 QY 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44  
 DB 287 CCAGTTAATAATGAAGTATTACGAACAATACACAACTTCAACCACTGGAG 346  
 QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaLeuIlePhe 59  
 DB 347 GAGGAATGACCACTGATGCTATTATTACAGTGGGATCGTCTGGCTGGTGGCT 406  
 QY 60 GlyTyrIleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrIysArgIleArg 79  
 DB 407 GCTTATCATCCAGGTTTCTTCTGGTGGCAGCAGGAGACAGATACTCAAAATTAGA 466  
 QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
 DB 467 AAACAATTTTTCATGCTATCATGCGACAGGAGATGGCTGGTTGACGTGCATGACGT 526  
 QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspIysIleSerAspGlyIleGlyAsp 118  
 DB 527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAATCAATAGAGAAATGGCGAC 586  
 QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
 DB 587 AAAATTGGAATGTTCTTCAATCAATAGCAACATTTTCCACCGGTTTATAGTGGGGTTT 646  
 QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
 DB 647 ACACGTGGTGGGAAGTAACCTTGTGATTTGGCCATCAGCCCTGTTCTGGCACTTCA 706  
 QY 159 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
 DB 707 GCGCCATCTGGCAAGATACATCTTCTTACTGATAAAGAACTCTTGGCCTATGCA 766  
 QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
 DB 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 826  
 QY 199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218  
 DB 827 GGACAAAGAAAGAACTTGAA----- 847  
 QY 219 TyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeu 238

DB 848 -----AGGTACACAAAAATTTA 865  
 QY 239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258  
 DB 866 GAAGAAGCTAAAGGAATTGGGATAAGAAAGTATCACGCCCAACATTTCTATGTGGCC 925  
 QY 259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrThrSerLeuIle 278  
 DB 926 GCTTCTTATTGATCATCATCATATGCTCTGGCTTTCTGGTATGGACCTCTCTGGTC 985  
 QY 279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298  
 DB 986 CTCTCCAGTGAA-----TATTCTATTGGCAAGTACTCACTGTCTCTTTCTGTATTA 1039  
 QY 299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAArg 318  
 DB 1040 ATGGGGCTTTTATGATTGGCAGGCATCCCAAGCATTTGAAGCATTTGCAACGGAAGA 1099  
 QY 319 GlyAlaAlaPheHisIlePheGlnValIleAspLysProSerIleAspAsnPheSer 338  
 DB 1100 GGAGCAGCTTATGAATCTTCAAGATAATGACATAAACCAAGCATTTGACAGCTATTG 1159  
 QY 339 ThrAlaGlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPhe 358  
 DB 1160 AAGAGTGGACATAAACACAGATAATATTAAAGGAAATTTGGAATTTCAAAATGTTCACTTC 1219  
 QY 359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378  
 DB 1220 AGTTACCTTCTCGAAAAGAAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTAGAGT 1279  
 QY 379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeu 398  
 DB 1280 GGCACAGAGTGGCTGTTGGAAACAGTGGCTGCGGAAGAGCACGCCGTGACGCTG 1339  
 QY 399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418  
 DB 1340 ATGCAGAGCTCTATGACCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399  
 QY 419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438  
 DB 1400 ACCATAAATGTAAAGGCATCTTCGGGAAATTTACTGGTGTGGTGGTGGAGCCTGTGTTG 1459  
 QY 439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGlu 458  
 DB 1460 TTTCACACCATGATGCTGAAACATTCCTATGCGCGGAAATGTCACCATGGATGAG 1519  
 QY 459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478  
 DB 1520 ATTGAGAAAGCTGTTAAGGAAGCCAAATGCTATGATTTTATCATGAACCTACCTAATAA 1579  
 QY 479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498  
 DB 1580 TTTCACACTCTGGTGGAGAGAGGGGCCAGCTGAGTGGTGGTGGACAGAAACAGAAATC 1639  
 QY 499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518  
 DB 1640 GCATTCCTGGGCCCTGTTTGCACCCCAAGATTTCTTCTGCTGATGAGGCAAGCTCA 1699  
 QY 519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
 DB 1700 GCTCTGACACTGAAAGTGAAGCAGTGGTTTCAAGTGGCCCTGGATAAG 1747  
 RESULT 8  
 ID AAF86128 standard; cDNA; 4195 BP.  
 XX AAF86128;  
 AC AAF86128;  
 XX 25-JUN-2001 (first entry)  
 DE Cynomologous monkey P-glycoprotein cDNA variant 1.  
 XX

KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;  
KW efflux pump; ss.  
XX  
OS Macaca fascicularis.  
XX  
FH Location/Qualifiers  
FT 100..3951  
FT /\*tag= a  
FT /product= "PGP"  
FT /note= "p-glycoprotein"  
FT misc\_feature 376..384  
FT /\*tag= b  
FT /note= "Insertion of 9 nucleotides relative to PGP  
FT allelic variant AAF86127"  
XX  
XX WO200123565-A1.  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-US26592.  
XX  
XX 28-SEP-1999; 99US-0156921.  
XX 12-OCT-1999; 99US-0158818.  
XX (GENT-) GENTEST CORP.  
XX  
XX Stocker PJ, Steimel-Crespi DT, Crespi CL;  
PI WPI; 2001-316136/33.  
DR P-P5DB; AAB81065.  
DR  
XX  
XX Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
PT (PGP) and homologous PGP polypeptides are useful for predicting  
PT bioavailability of compound and increasing PGP transporter activity in  
PT cell -  
XX  
XX Example 1; Page 59-65; 84pp; English.  
XX  
XX This invention relates to a polynucleotide sequence encoding a  
CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
CC exporting small molecules across the cell membrane. The invention  
CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
CC and protein, and also that of an allelic variant. The PGP polynucleotide  
CC sequence is useful for increasing PGP transporter activity in a cell.  
CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
CC activity in a mammalian cell. They may also be used for increasing the  
CC bioavailability of a drug. The present sequence represents cDNA encoding  
CC cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide  
CC insert compared to the PGP allelic variant given in AAF86127.  
XX  
XX Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;  
SQ  
  
Alignment Scores:  
Pred. No.: Length: 1,99e-150 4195  
Score: 1557.00 Matches: 306  
Percent Similarity: 72.17% Conservatives: 96  
Best Local Similarity: 54.94% Mismatches: 103  
Query Match: 56.85% Indels: 52  
DB: 22 Gaps: 5  
  
US-09-873-409-7 (1-541) x AAF86128 (1-4195)  
  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 250 ATGGTGGTGGAACTTTGGTGGCCATCATCGAGCTGGAGTCTCTCTCATGATGCTG 309  
QY 21 ValLeuGlyIleLeuMetSerAspAsnLeuIleSer-----GlyCys 33  
DB 310 GTGTTTGGAGACATGACCGATACCTTTGCAAAATGCGAGAAATTTAGGAGATTAGGAGCT 369  
QY 34 LeuValGlnThrAsnThrTyrSer----- 41

DB 370 CTGTTGTTTAAACAACACACTAATAGCAGTAATATCACTGATACAGTCCCGCTCAATGATCTG 429  
QY 42 -----PhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIle 58  
DB 430 GAGGAGATATGACACAGGTATGCCCTATTATTACAGTGGAAATGGTGGTGGTGGTGGT 489  
QY 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78  
DB 490 GCTGCTTACATTACAGTTCTATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 549  
QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98  
DB 550 AGAAACAACAGTTTTTTCATGCTATATGACACAGGAGATAGGCTGGTTTATGTGACAGAT 609  
QY 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117  
DB 610 GTTGGGAGCTTAAACACCGGCTTACAGATGATGCTCCAGATTAATGAAGGAATGGT 669  
QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137  
DB 670 GACAAATTTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGCA 729  
QY 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157  
DB 730 TTTACACGTGGTGGAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTG 789  
QY 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177  
DB 790 TCAGTGTGAGCTCTGGCAAGATAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 849  
QY 178 SerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPhe 197  
DB 850 GCAAAAGCTGGAGCAGTAGTCTCAAGAGGTCTTGGCAGCAATTAGAATCTGATTTGATTT 909  
QY 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217  
DB 910 GGAGGACAAAGAAAGAACTCGAA----- 933  
QY 218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu\*\*PheValArgTyrThrGlnAsn 237  
DB 934 -----AGGTAACAACAAAT 948  
QY 238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257  
DB 949 TTAGAAGAGCTAAAGAAATTTGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGT 1008  
QY 258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277  
DB 1009 GCTGCTTTCTCTCTTATCTATCATCTTATGCTCTGCTCTGCTCTGCTATGGGACCACTTG 1068  
QY 278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297  
DB 1069 GTCCCTCTCAAGGAA-----TATCTATGACAAAGTACTCACTGATTTCTTTCTGTA 1122  
QY 298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317  
DB 1123 TTAATTTGGGCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCAATTTGCAATGCA 1182  
QY 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysIleAspSerIleAspAsnPhe 337  
DB 1183 AGAGGAGCAGCTTTTGAATCTTCAAGATAATTTGAATAAAGCAAGTATTGACACTAT 1242  
QY 338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357  
DB 1243 TCGAGAGTGGGCACAAACAGATATATTAAAGGAAATTTGGAATTCAGAAATGTTCCAC 1302  
QY 358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377  
DB 1303 TTCAGTTACCAATCTCGAAAGAGTATTAGATCTTGAAGGCTTGAACCTGAGGTCGAG 1362  
QY 378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397

Db 1363 AGTGGCAGACGGTCCCTGGTTGGAAACAGCGGCTGTGGGAAGACCAACGGTCCAG 1422  
 QY 398 LeuLeuGluArgLeuTyrAspProAspAspGlyPheLeuMetValAspGluAsnAspIle 417  
 Db 1423 CTGATCGAGAGGCTTTATGACCCACAGAGGGCATGTCAGTGTTCATGGACAGGATATT 1482  
 QY 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437  
 Db 1483 AGGACCAATAACGTAAGTTTCTACGGAAATCATCGGTGTGTGAGTCAGGAACCTGTA 1542  
 QY 438 LeuPheGlyThrThrLeuSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457  
 Db 1543 TTGTTTCCACACCATAGTACGTAAGCAATTCGCTATGCTGTGCTGAGCATGCCATGAT 1602  
 QY 458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477  
 Db 1603 GAGATTGAGAAAGCTGTCAAGGAAGCCAAATGCTATGACTTTATCATGAACCTGCCCTCAG 1662  
 QY 478 LysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArg 497  
 Db 1663 AAATTTCACACCTGGTTGGAGAGAGGGGCCCGCTGAGTGTGGCGAGAGCAGAGG 1722  
 QY 498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517  
 Db 1723 ATCGCCATTGCAGTGCCTCTGGTTGCGAACCCCAAGATCCTCTGCTGGACGAGGCCAG 1782  
 QY 518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
 Db 1783 TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1833  
 RESULT 9  
 AAZ49332  
 ID AAZ49332 standard; cDNA; 3860 BP.  
 XX AC AAZ49332;  
 XX DT 14-MAR-2000 (first entry)  
 XX DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.  
 XX KW Multidrug resistance; MDR-1; P-glycoprotein;  
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassaemia;  
 KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
 KW cytokine; wild-type; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX CDS 1..3843  
 FT /tag= a  
 FT /product= "Human wild-type MDR-1 protein"  
 FT mutation replace (553..555, GTT)  
 FT /tag= b  
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given  
 in AAZ49333"  
 XX MO9961589-A2.  
 XX PD 02-DEC-1999.  
 XX 27-MAY-1999; 99WO-US11825.  
 XX 28-MAY-1998; 98US-0086988.  
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Sorrentino B, Bunting K;  
 XX WPI; 2000-072615/06.  
 XX P-PSDB; AAY58186.  
 XX DR

PT Ex vivo expansion of hematopoietic stem cells transduced with a  
 PT sequence encoding human multidrug resistance-1, used for bone marrow  
 PT transplantation -  
 PS Claim 10; Page 68-70; 113pp; English.  
 XX This sequence represents cDNA encoding human wild-type  
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane  
 CC efflux pump, responsible for the export of drugs from cells,  
 CC particularly cancer cells. Wild-type MDR-1 shows increased  
 CC resistance to etoposide and decreased resistance to vinca  
 CC alkaloids compared with a mutant form (AAY58187) where the Gly at  
 CC position 185 is replaced by Val. The invention relates to transducing  
 CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified haematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC haematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from haematopoietic stem cells, e.g., thalassaemia,  
 CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in haematopoietic stem cells.  
 CC Haematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC haematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.  
 XX SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,52e-150 Length: 3860  
 Score: 1555.50 Matches: 306  
 Percent Similarity: 73.23% Conservative: 100  
 Best Local Similarity: 55.23% Mismatches: 99  
 Query Match: 56.79% Indels: 49  
 DB: 21 Gaps: 6  
 US-09-873-409-7 (1-541) x AAZ49332 (1-3860)  
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 151 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGTGGACTTCCTCTCATGATGCTG 210  
 QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 211 GTGTTGGAGAAATGACAGATATCTTTGCCAATGCGAGGAATTTAGAGATCTGATGTCA 270  
 QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 271 AACATCACTAATAGAGATGATATCATGATACAGGGTCTTTCATGATCTGGAGGAAGAC 330  
 QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 331 ATGACACAGATATGCTTATTATTACAGTGAATTTGGTGTGGGGTGTGCTGTCTGCTTAC 390  
 QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGGAAGCAAAATACACAAAATTAGAAACAG 450  
 QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
 Db 451 TTTTTCATGCTATATAGCAGAGATAGGCTGTTGATGTGCACGATGTTGGGAG 510  
 QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 511 CTTTAACACCCGACTTACAGATGATGCTCTAGATTAAATGAAGGTATTGGTGACAAAT 570  
 QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 571 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGT 630

```
QY 141 GlyTyrLysLeuThrValThrLeuSerThrSerProLeuLeuMetAlaSerAlaAla 160
DB 631 GGTGGAGCTAACCTTCTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTGAGCTGCT 690
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 691 GTCTGGCAAGATACATCTATCTTCTATCTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 751 GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGAATGTGATTCATTTGGAGGACAA 810
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAenIleThrArgTyrAlaTrpPheTyrPhe 220
DB 811 AAGAAGAACTTGAA----- 825
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAenLeuLysAsp 240
DB 826 -----AGGTACCAACAAATAATTTAGAAAGAA 849
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 850 GCTAAAGAAATTGGGATAAAGAAAGCTATTACAGCCAAATATTTCTATAGTGTGCTTTC 909
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
DB 910 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGCTATGGGACCACCTTGGTCTCTCA 969
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 970 GGGGAA-----TATTCTATTGGCAAGTACTCACTGTATTCTTTCTGTATTAAATTGGG 1023
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1024 GCTTTAGTTGGACAGCACTCCAGCAATTGAAGCAATTTGCCAATTCGAAGAGAGCA 1083
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1084 GCTTATGAATCTTCAAGATAATTGATAATAGCCAAATATTGACAGCTATTGCAAGAGT 1143
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1144 GGGCACAACCAAGATAATATTAGGGAAATTTGGAAATTCAGAAATGTTCACTTTCAGTTAC 1203
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1204 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGGCTGAACTGAAGGTGCAGATGGGCGAG 1263
QY 381 ThrValAlaLeuValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGGGAAGAGCACACAGTCCAGCTGATGCGAG 1323
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1324 AGGCTCTATGACCCACAGAGGGGATGGTTCAGTGTGGTGGACAGGATATAGGACCATA 1383
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1384 AATGTAAAGTTTCTACGGGAATCATTTGGTGTGGTGCAGTACAGAACTGTATTGTTGGC 1443
QY 441 ThrThrIleSerAsnAenIleLysTyrGlyArgAspAspValThrAspGluMetGlu 460
DB 1444 ACCACATAGCTGAAGCAATCTGCTATGSCCGTGAANAATGTACCATGGATGAGATTGAG 1503
QY 461 ArgAlaAlaArgGluAlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAsn 480
DB 1504 AAGCTGTCAAGGAAGCAATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 1563
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1564 ACCCTGGTTGGAGAGAGGGGCCCATGTTGAGTGGTGGGCGAAGAGGAGGATCGCCATT 1623
```

```
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1624 GCACGTGCCCTGGTTGGCAACCCCAAGATCTCTGCTGGATGAGCCACGTCAGCCTTG 1683
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1684 GACACAGAAGCGAAGCAGTGGTTGAGTGGTCTGGATAAG 1725
RESULT 10
ABA94365
ID ABA94365 standard; DNA; 3860 BP.
XX ABA94365;
AC ABA94365;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human BCRP DNA related seq Id No. 1.
XX
KW Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW hepatotropic; neurodegenerative; cytotstatic; antianemic; muscular; BCRP;
KW cardiac; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3843
FT FT /*tag= a
XX
PN WO200192877-A2.
XX
PD 06-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US17459.
XX
PR 31-MAY-2000; 2000US-0584586.
PR 29-MAY-2001; 2001US-0866866.
XX
(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Sorrentino B, Schuetz J;
DR WPI: 2002-114368/15.
DR P-PSDB; ABB07266.
XX
PT Identifying a stem cell, for treating e.g., muscular dystrophy,
PT myocardial infarction, Parkinson's disease, or neurodegenerative
PT disorders, comprises detecting the expression of an ATP transport
PT protein (BCRP) by a cell
XX
PS Disclosure; Page 53-55; 87pp; English.
XX
CC The invention provides a method of identifying and/or isolating a stem
CC cell that involves detecting the expression of an ATP transport protein
CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC comprising stem cells. The isolated stem cells may be used in the
CC treatment of diseases such as muscular dystrophy, degenerative liver
CC disorder, myocardial infarction, Parkinson's disease, degenerative
CC disorders of the brain, and for tissue regeneration or replacement.
CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC treatment of leukemia) and for ex vivo gene therapy for treating blood
CC diseases such as sickle cell anemia and thalassemia. The stem cells can
CC also be used as cell targets in gene therapy protocols. The present
CC sequence represents a sequence related to the BCRP for which no relevant
CC information has been provided in the specification.
XX
SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 2,528-150 Length: 3860
Score: 1555.50 Matches: 306
Percent Similarity: 73.29% Conservative: 100
Best local Similarity: 55.23% Mismatches: 99
Query Match: 56.79% Indels: 49
```

DB: 24 Gaps: 6

US-09-873-409-7 (1-541) x ABA94365 (1-3860)

Qy 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 151 ATGGTGGGAACCTTGGCTGCCATCCATGGGCTGGACCTTCCTCTCATGATGCTG 210  
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCCAATGCAGGAAATTTAGAGATCTGATGCA 270  
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTCTTTCATGAATCTGCAGGAAGAC 330  
 Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 331 ATGACACAGATATGCGCTATTATTACAGTGAATTTGGTCTGGGGTCTGTGCTGCTTAC 390  
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 391 ATTCAAGTTTCATTTGGTCCCTGGCAGCTGGAAGACAAATACACAAATTTAGAAAAACAG 450  
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
 Db 451 TTTTTCATGCTATAATGCCACAGGAGTAGCTGCTTTTGATGTCGACGATGTTGGGGAG 510  
 Qy 102 LeuAsnThrArgMetThr---AspIleAspIleValIleSerAspGlyIleGlyAspLysIle 120  
 Db 511 CTTAAACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGGTATTGGTGACAAAATT 570  
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 571 GGAATGTTCTTTGAGTCAATGGCAACATTTTCACCTGGGTTTATAGTAGGATTACAGT 630  
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleValAlaSerAlaAla 160  
 Db 631 GGTGGGAAGCTAAACCTTGATGATTTGGCCATCAGTCTGTTCTTGGACCTGTGAGTCT 690  
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 691 GTCTGGCCAAAGATACATCTTCATTTACTGATAAAGAACTCTTACGCGTATGCCAAAAGCT 750  
 Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 751 GGACAGTAGCTGAGAGAGTCTGGCAGCAGCATTTAGAACTGTGATTCATTTGGAGACAA 810  
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
 Db 811 AAGAAAGAACTTGAA----- 825  
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
 Db 826 -----AGGTACAACAAATAATTTAGAGAA 849  
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
 Db 850 GCTAAAAGAAATGGGATPAAAGAAAGCTATTACGCCAATATTTCTATAGGTGCTGCTTTC 909  
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
 Db 910 CTGCTGATCATGTCATCTTATGCTCTGGCCCTCTGGTATGGGACCACTTGGTCTCTCA 969  
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 Db 970 GGGGAA-----TATTCTATTGGCAAGTACTCACCTGATCTCTTTCTGTATTAAATGGG 1023  
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
 Db 1024 GCTTTTGTGTGGCAGGCGATCTCCAAGCATTTGCAAAATTCAGAGAGAGCA 1083  
 Qy 321 AlaPheHisIlePheGlnValIleAspGlyLysProSerIleAspAsnPheSerThrAla 340

XX Isolated polynucleotide encoding a rat mdrib2 drug transporter  
PT polypeptide, useful for treatment of e.g. cancer, autoimmune disease,  
PT central nervous system disorders -

XX Claim 7; Column 13-17; 17pp; English.

XX This sequence represents a polynucleotide encoding a 1275 amino acid  
CC rat mdrib2 multispecific drug transporter polypeptide. The mdrib2  
CC nucleotide sequence has cytostatic, antiinflammatory, cardiant,  
CC neuroprotective, immunosuppressive and nephrotropic activity.  
CC Understanding the functioning of the mdrib2 polynucleotide and protein in  
CC transgenic animal models is useful for treating and preventing diseases  
CC such as cancer, inflammation, cardiovascular disease, central nervous  
CC system disorders, autoimmune disease, and kidney disease. The use of the  
CC protein in cell based, membrane based, or binding assays may enhance drug  
CC formulation, selection of formulation excipients and compound design.

XX Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

#### Alignment Scores:

Pred. No.: 5,88e-150 Length: 4233  
Score: 1552.50 Matches: 308  
Percent Similarity: 72.20% Conservative: 92  
Best Local Similarity: 55.60% Mismatches: 105  
Query Match: 56.68% Indels: 49  
DB: 21 Gaps: 5

US-09-873-409-7 (1-541) x AAZ90198 (1-4233)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 170 ATGGCTCTGGGAACCTCGCTGCTATCATCCAGGAACTCGTCCCTCTCTGATGCTG 229  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
Db 230 GTGTTCGGATACATGACATAGATTTTACCACAGAGAGACCCGCATTCGCCGAGCGTT 289  
Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTySerPheArgLeu----- 45  
Db 290 ACTAATCAAGTGAATCAACAGTACACAGCGCTCAGCGACAGCATCTGGAGGAGGAC 349  
Qy 46 -----ThrLeuTyTyTyValGlyIleGlyValAlaAlaLeuIlePheGlyTy 61  
Db 350 ATGGCCATGTAGCCCTACTATTACACGGCATTTGGTGGCGGTGCTCATCTGTTGCTAC 409  
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleArgIleGln 81  
Db 410 ATCCAGGTTTCACTTTGGTGCCTGGCGAGCTGGGAGACAAATACACAGATTAGGCGAGAG 469  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyPheAspSerCysAspIleGlyGlu 101  
Db 470 TTTTTCATGCCATCAATCAATCAGGAGATAGGCTGTTTGACGTGAATGACGCTGGGAG 529  
Qy 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120  
Db 530 CTCAACACCCGGCTCACAGATGACGCTCCAAAATAATGACGGAATTGGTGACAACTT 589  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 590 GGAATGTTCTTTCAGTCCCAACAGCATTTTTCAGCCGGTTTATAATAGGATTATAAGT 649  
Qy 141 GlyTyPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 650 GGTGGAAGCTAACCTTGTAATTTTGGCGGTGAGCCCTCTTATTGGGTGTCATCTGCC 709  
Qy 161 AlaCysSerArgMetValIleSerLeuSerIleGlyLeuSerAlaTySerLysAla 180  
Db 710 ATGTGGGCAAGGTACTGACTTCACTTACTAATAAGGAACCTCCAGGCTTATGGAAGCT 769  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 770 GGAGCAGTTGCCGAAGAGTCTTAGCAGCCATCAGAACCTGTGATTGCGTTGGAGGACAA 829

Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyAlaTrpPheTyPhe 220  
Db 830 AAGAAGGAACTTGAA----- 844  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyThrGlnAsnLeuLysAsp 240  
Db 845 -----AGTCAATATAAAATTTAGAGAA 868  
Qy 241 AlaLysAspPheGlyIleLeuArgThrIleAlaSerLysValSerLeuGlyAlaValTy 260  
Db 869 GCTAAAGAGTTGGCTAATAAGAACCCATCAGGCCAACATTTCCATAGTATTGCTCTAC 928  
Qy 261 PhePheMetAsnGlyTyTyGlyLeuAlaPheTyTyGlyThrSerLeuIleLeuAsn 280  
Db 929 CTGTGGTCTATGCTCTATGCACTGGCATTCGTGTATGGACCTCCTGTGCTCTCTCA 988  
Qy 281 GlyGluProGlyTyThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 989 AATGAA-----TATTCTATTGGCAAGTGTCTTACCGTCTTCTTCTATTTATTGGGG 1042  
Qy 301 SerTyTyCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1043 ACTTTCAGTATTGGACATTTAGCCCCCAACATAGAGCCCTTTTGCAATGCAAGAGGGCA 1102  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1103 GCCTATGAAATCTTCAAGATAATTGATAATAGCAAGCAATTCACAGCTTCTCAACCAAG 1162  
Qy 341 GlyTyTyLysProGluSerIleGlyThrValGluPheLysAsnValSerPheLeuTy 360  
Db 1163 GGACACAAACACAGACAGTATAATGGGAATTTTGGAAATTTTAAATAATGTTTACTTCAACTAC 1222  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1223 CCATCAGAGTGAAGTTAAGATCTTGAAGGCCCTCACTGAAGTGAAGAGCGGCGAG 1282  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1283 ACGGTAGCCCTGTTGGCAACAGTGGCTGTGGGAAAGCACAACTGTCCAGCTGTCTGCAG 1342  
Qy 401 ArgLeuTyTyAspProAspAspGlyPheIleMetValAspLysAsnAspIleArgAlaLeu 420  
Db 1343 AGGCTCTACAGCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGCCATC 1402  
Qy 421 AsnValArgHisTyArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1403 AATGTGAGGTATCTCGGGGAATCATTTGGGTGGTGGTGTGAGTCAGGAACCCGTGCTTGTCC 1462  
Qy 441 ThrThrIleSerAsnAsnIleLysTyTyArgAspAspValThrAspGluGluMetGlu 460  
Db 1463 ACCAGATTGCCGAAACATTCGTATGCGCGAGAAACGTCACCATGGATGAGATAGAG 1522  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyArgPheIleMetGluPheProAsnLysPheAsn 480  
Db 1523 AAAGCTGTCAAGGAAGCCCAATGCTATGATCTTCATCATGAAACTGCCCCACAAATTAAC 1582  
Qy 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1583 ACCCTGTTGGTGGAGAGGGCGGCGAGCTGAGTGGGGGACAGAAACAGAGGATCGGCATT 1642  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1643 GCCCGGCCCTGTGTCGCAACCCCAAGATCTTTTGTGGATGAGGCCACGTCAGCCTTG 1702  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
Db 1703 GACACAGAAAGGAGCGGTTCAGGCGGCTCTGGATAAG 1744  
RESULT 12  
AAF27498  
ID AAF27498 standard; cDNA; 4233 BP.  
XX





```

Db 1223 CCATCAGAGTGAAGTCTTGAAGGCGCTCAACCTGAAGTGAAGAGCGGCGAG 1282
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
Db 1283 ACGGTAGCCCTGGTGGCAACAGTGGCTGGGAAAGACACAACCTGCCAGCTGCTGCAG 1342
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db 1343 AGGCTCTACGACCCCATAGAGGGCGAGGTTCAGTATCGACGAGCAGACATCAGGACCATC 1402
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db 1403 AATGTAGGTATCTCGGGAATCATTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1462
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db 1463 ACCAGATTGCCGAAACATTCGCTATGCGCGAGAAACGTCACCATGGATGAGTAGAG 1522
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db 1523 AAAGCTGTCAAGGAAGCCCAATGCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582
Qy 481 ThrValGlyGlyGlyGlyAlaGlnMetSerGlyGlyGlyGlyGlnArgIleAlaIle 500
Db 1583 ACCCTGGTGGTGGAGAGGGGGCGAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 1643 GCCCGGCCCTGGTCCGCAACCCCAAGATCTCTTTGTTGGATGAGGCCACGTCAGCCTTG 1702
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 1703 GACACAGAAAGCGAGCCGCTGGTTTCAGGCGCTCTCGATAAG 1744

RESULT 13
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX
XX WO200123540-A2.
XX
PD 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
XX WPI; 2001-235373/24.
XX
DR P-PSDB; AAE00303.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -

```

```

XX Claim 3; Page 58-63; 11lpp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein, PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.: 5.97e-150 Length: 4279
Score: 1552.50 Matches: 302
Percent Similarity: 71.94% Conservative: 98
Best Local Similarity: 54.32% Mismatches: 105
Query Match: 56.68% Indels: 51
Db: 22 Gaps: 5

US-09-873-409-7 (1-541) x AAD03488 (1-4279)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTGGTGGGACAAATGGCTGCCATCATCGATGAGCTGCACTCCCTCTCATGATGCTG 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCCAATGCGAGGAATTTCAAGAAACAAACTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44
Db 287 CCAGTTATAATTAATGAAGATATTAGCAACAATACACAACTTTCATCAACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuPhe 59
Db 347 GAGGAAATGACCACGATATGCCATTATTATACGTGGATCGGTGCTGGCTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGTTTCATTCTGTCCTGGCAGCAGGAGAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 467 AAACAATTTTTCATGCTATCATCGACAGAGATGGCTGGTTTCACGTCGATGACGTT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGGAGCTTAAACACCCGGCTCACAGCATGTCTCAAAATCAATGAAGAAATTCGGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAATTTGGAATGTTCTTTTCATCAATAGCAACATTTTTCACCGTTTATAGTGGGTTT 646
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTTGGAAGTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTCA 706
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCCGCCATCTGGGCAAGATACTACTTCTTCACTTACTACTATAAGAACTCTTGGCCTATGCA 766

```



Qy	359	AsnTyrProSerhrgProSerlleLeuLysGlyLeuAsnLeuArglleLysSer	378
Db	1270	AGTTACCTCTCTCGAAGAAGATTAGATCTTTAAAGGGTCTCAACTCGAAGGTTTCAGAGT	1329
Qy	379	GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu	398
Db	1330	GGGACAGACAGTGGCGTGGTGGGACAGTGGCTCGCGGAAGAGCAGCACCGTCGACGCTG	1389
Qy	399	LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg	418
Db	1390	ATGCAGAGGCTCTATGACCCACAGATGCATGGTCTGTATTGTGGACAGGACATTAGG	1449
Qy	419	AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValValSerGlnGluProValLeu	438
Db	1450	ACCATAAATTAAGGCATCTTCGGGAATAATTACTGGTGTGGTGAGTCAGAGCCGTGTGTG	1509
Qy	439	PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu	458
Db	1510	TTTGCCACACAGATAGCTGAAACATTCGCTATGATGCCGGAATAATGTCAACATGATGAG	1569
Qy	459	MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys	478
Db	1570	ATTGAGAAAGCTGTTAAGGAAGCCAATGCTATGATTTTATCATGAAACTACCTAATAAA	1629
Qy	479	PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle	498
Db	1630	TTTGACACTCTGGTGGAGAGAGAGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATC	1689
Qy	499	AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer	518
Db	1690	GCCATTGCTGGGCGCTGGTTCGCAACCCCAAGATCTCTCTGCTGATGAGGCAACGTC	1749
Qy	519	AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys	534
Db	1750	GCTCTGGACACTGAAAGTGAAGCAGTGGTTTCAGGTGGCCCTGGATAAG	1797
RESULT 15			
AAV32645			
ID	AAV32645 standard; cDNA; 4669 BP.		
AC	AAV32645;		
XX			
DT	23-SEP-1998 (first entry)		
DE	Human P glycoprotein (Pgp) cDNA.		
XX			
KW	Human P glycoprotein; Pgp; multi-drug resistance; cancer;		
XX	UIC2 monoclonal antibody; mAb; cytotoxic; transmembrane efflux pump; s		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	5'UTR	1..424	
FT		/*tag= a	
FT	CDS	425..4267	
FT		/*tag= b	
FT		/product= "wild-type Pgp"	
FT	3'UTR	4268..4469	
FT		/*tag= c	
XX			
PN	W09821325-A1.		
XX			
PD	22-MAY-1998.		
XX			
PF	17-NOV-1997; 97WO-US21214.		
XX			
PR	15-NOV-1996; 96US-0752447.		
XX			
PA	(INGE-) INGENEX INC.		
PA	(UNII) UNIV ILLINOIS FOUND.		
PI	Mechetner E, Roninson IB;		

XX WPI: 1998-297930/26.  
DR P-PSDB; AAW48997.  
XX  
XX Immunological reagent specific for P-glyco:protein - useful for  
PT detecting multi-drug resistant cancer, isolating haematopoietic  
PT cells and selective cell killing  
XX  
XX  
PS Claim 1; Fig 1A; 89pp; English.  
XX  
XX The present sequence represents the human P glycoprotein (Pgp)  
CC CDNA which encodes the wild-type Pgp protein. Pgp is a transmembrane  
CC efflux pump protein involved in multi-drug resistance of cancer cells.  
CC The invention provides methods for developing and using immunological  
CC reagents specific for certain mutant forms of Pgp and wild-type Pgp in  
CC a conformation associated with substrate binding or in the presence  
CC of ATP depleting agents. An example of the immunological reagent  
CC is the UIC2 monoclonal antibody (mAb). mAb UIC2 specifically binds  
CC to Pgp in a particular biochemical conformation and is capable of  
CC inhibiting drug efflux from Pgp-expressing cells. The immunological  
CC reagents are claimed to be useful for detecting Pgp expression in  
CC mammalian cells, including low level expression, particularly in cancer  
CC cells to diagnose multi-drug resistance. The invention claims that  
CC these immunological reagents are more specific than known reagents for  
CC detecting Pgp and they also eliminate the need for costly and laborious  
CC screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.  
XX  
SQ Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

Alignment Scores:  
Pred. No.: 6.77e-150 Length: 4669  
Score: 1552.50 Matches: 305  
Percent Similarity: 73.29% Conservative: 101  
Best Local Similarity: 55.05% Mismatches: 99  
Query Match: 56.68% Indels: 49  
DB: 19 Gaps: 6

US-09-873-409-7 (1-541) x AAV32645 (1-4669)  
QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 575 ATGGTGGGGAACCTTTGGTGGCCATCATCCATGGGGCTGGAGTCTCTCATGATGCTG 634  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 635 GTGTTGGAGAAATGACAGATATCTTGCRAATGCAGGAATTTAGAGATCTGATGCA 694  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 695 AACATCAATAAGATGATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754  
QY 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 755 ATGACCAAGTATGCTATTATTACAGTGAATGGTGGGGTGGTGGTGGTGGTGGTGGTTC 814  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 815 ATTCAAGTTTCATTGTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 875 TTTTTCATGCTATTATCGCAGCAGGAGATAGCTGGTTGTATGTGCACGATGTTGGGAG 934  
QY 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120  
DB 935 CTTAACACCCGACTTACAGATGATGCTCCCAAGATTAATGAAGGAATGGTGACAAAT 994  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTCAGTGGTTTATAGTAGGATTTACAGT 1054  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160

DB 1055 GGTGGAAAGCTAACCTTGTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGCTGCT 1114  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysLeuSerLysLeuSerLysLeuSerLysAla 180  
DB 1115 GTCTGGGCAAGATGATCTATCTTCACTTACTGTATAAAGAACTCTTAGCGTATGCAAAAGCT 1174  
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGACAA 1234  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 1235 AAGAAAGAACTTGA----- 1249  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 1250 -----AGGTACAACAACAAATTTAGAGAA 1273  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 1274 GCTAAAAGAAATTTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTC 1333  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrThrSerLeuIleLeuAsn 280  
DB 1334 CTGCTGATCTATGCACTTATGCTGGCCCTTCTGGTATGGACACCTTGGCTCTCA 1393  
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 1394 GGGGAA-----TATTTCTATGGCAAGTACTCACTGTATTTCTTTCTATTAATTTGGG 1447  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla9GlyAla 320  
DB 1448 GCTTTTAGTGTGGACAGGCATCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGACGA 1507  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 1508 GCTTATGAATCTTCAAGATAATTGATAATGACCAAGTATTGACAGCTATTGCAAGAGT 1567  
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 1568 GGGCACAAACCATGATAATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTAC 1627  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 1628 CCATCTCGAAAAGAAAGTTAAGATCTTGAAGGGCTTGAACCTGAAGGTGCAGAGTGGGCAG 1687  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 1688 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCGAG 1747  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
DB 1748 AGGCTCTATGACCCACAGAGGGGATGGTGCAGTGTTCATGGACAGGATATTAGCACATA 1807  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 1808 AATGTAAGTTCCTACGGGAAATCATTTGGTGTGGTGTGAGTCAGGAACCTGATTTGTC 1867  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
DB 1868 ACCACGATGCTGAAAACATTCGCTATGGCCGTGAAATGTCCACATGATGAGATTTGAG 1927  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 1928 AAGCTCTCAAGNAGCAATGCTATGATCTTTATCATGAACTGCCTCATAAATTTGAC 1987  
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 1988 ACCCTGTTGGAGAGAGAGGGGCCAGTGTGAGTGTGGGAGAGAGAGAGATGCCATT 2047  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 2048 GCACGTGCCCTGGTTGCAACCCCAAGATCTCTGCTGATGAGGCCACGCTCAGCCTTG 2107

Qy 521 AspSerGluSerIysSerAlaValGlnAlaAlaLeuGluLys 534  
|||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 2108 GACACAGAAAGCGACGCTGGTTTCAGGTGGCTCTGGATAAG 2149

Search completed: March 30, 2003, 03:21:35  
Job time : 587.737 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 3721.1 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-7  
Perfect score: 2739  
Sequence: 1 MILGILASLVNGACLPMLP.....SESKSAVQAALKEKTPRYSP 541

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115420\_19240/app\_query.fasta\_1.7544  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 30544 @runat\_27032003\_115420\_19240 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_estcl:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1092	39.9	669	13	BM013981	BM013981 603639619
2	1068.5	39.0	1033	14	BM926413	BM926413 AGENCOURT
3	1029	37.6	760	13	BM016204	BM016204 603642659
4	918	33.5	545	9	AL040762	AL040762 DKFP2434C
5	903	33.0	929	14	BQ882401	BQ882401 AGENCOURT
6	856.5	31.3	948	13	BM471690	BM471690 AGENCOURT
7	750	27.4	894	12	BF584668	BF584668 602098406
8	744.5	27.2	872	11	AK020318	AK020318 Mus muscu
9	689.5	25.2	2676	11	AK014319	AK014319 Mus muscu
10	659	24.1	943	9	AL520322	AL520322 AL520322
11	654	23.9	405	9	AA243820	AA243820 zrf7906.r
12	650	23.7	550	13	BJ075336	BJ075336 BJ075336
13	648	23.7	803	12	BG298756	BG298756 602396681
14	637	23.3	780	12	BG587938	BG587938 EST489713
15	625	22.8	559	12	BG692596	BG692596 602248949
16	621	22.7	606	14	BQ262340	BQ262340 fz85a08.x
17	621	22.7	681	14	BQ06064	BQ06064 QGG9N16.Y
18	604.5	22.1	932	17	AZ670821	AZ670821 ENTJUN69TF
19	602.5	22.0	913	17	BH155700	BH155700 ENTROS54TR
20	599	21.9	880	17	AZ687805	AZ687805 ENTJUS2TF
21	596.5	21.8	871	17	AZ682350	AZ682350 ENTXB16TF
22	590.5	21.6	891	17	AZ679807	AZ679807 ENTHI67TR
23	585.5	21.4	853	17	AZ679807	AZ679807 ENTHI67TR
24	583	21.3	2564	11	AK020022	AK020022 Mus muscu
25	582.5	21.3	787	12	BG584063	BG584063 EST485823
26	581	21.2	586	13	BJ336461	BJ336461 BJ336461
27	579	21.1	661	12	BG302675	BG302675 fl36b04.x
28	578	21.1	575	10	AK342644	AK342644 fl81f06.x
29	577.5	21.1	886	17	AZ540627	AZ540627 ENTQ18TF
30	577	21.1	835	13	BJ349604	BJ349604 BJ349604
31	576.5	21.0	947	17	AZ683753	AZ683753 ENTJL96TF
32	576	21.0	547	13	BJ333166	BJ333166 BJ333166
33	574.5	21.0	885	17	AZ686798	AZ686798 ENTZF28TR
34	574	21.0	542	13	BJ339421	BJ339421 BJ339421
35	572.5	20.9	860	17	AZ543293	AZ543293 ENTQ80TF
36	572.5	20.9	886	17	BH139685	BH139685 ENTNA47TR
37	571	20.8	715	14	BQ869512	BQ869512 QGD6G09.Y
38	570	20.8	814	13	BJ352204	BJ352204 BJ352204
39	570	20.8	820	13	BJ356540	BJ356540 BJ356540
40	570	20.8	1019	12	BG248052	BG248052 602359987
41	569	20.8	1341	11	AY108285	AY108285 Zea mays
42	565	20.6	657	14	BQ802058	BQ802058 WHE2821.F
43	564	20.6	636	13	BJ386538	BJ386538 BJ386538
44	564	20.6	750	13	BJ445883	BJ445883 BJ445883
45	562	20.5	1813	11	U66688	U66688 Homo sapien

ALIGNMENTS

RESULT 1  
BM013981  
LOCUS 603639619P1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5415556 5',  
DEFINITION mRNA sequence.  
ACCESSION BM013981  
VERSION BM013981.1 GI:16528335  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (Bases 1 to 869)  
AUTHORS NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12060 row: f column: 05  
High quality sequence stop: 664.  
Location/Qualifiers

## FEATURES

source

1..669  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5415556"  
/clone\_lib="NIH MGC 87"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC library."  
BASE COUNT 203 a 122 c 162 g 182 t

## ORIGIN

Alignment Scores:  
Pred. No.: 2,18e-123 Length: 669  
Score: 1092.00 Matches: 215  
Percent Similarity: 99.54% Conservative: 2  
Best Local Similarity: 98.62% Mismatches: 0  
Query Match: 39.87% Indels: 1  
DB: 13 Gaps: 0

US-09-873-409-7 (1-541) x BM013981 (1-669)

QY 292 AlavalPheSerValIleHisSerSerTyrCysIleGlyAlaValProHisPhe 311

DB 17 AGTGTTTCTTAATGTAATCCATAGCAGTATTGCGAGCAGAGTCCCTCACTTT 76

QY 312 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 331

DB 77 GAACCTTCGCATAGCCGAGGAGTGCCTTTCATATTTTCCAGGTATTGATAGAAA 136

QY 332 ProSerIleAspPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 351

DB 137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAACTGTG 196

QY 352 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGly 371

DB 197 GAATTTTAAAAATGTTCTTTCAATATATCCATCAAGACCATCTCAAGATTCGAAAGGT 256

QY 372 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 391

DB 257 CTGAATCTCAGATTAGTCTGGAGACACAGTCGCTTGTGCTCATGCGAGTGGG 316

QY 392 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 411

DB 317 AAGAGTACGGTAGTCCAGCTTCTCGACAGGTATATGATCCGATGATGCTTATCATG 376

QY 412 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 431

DB 377 GTGGATGAGAATGACATCAGAGCTTTTAAATGTGGGCATTTATCGAGACCATATTGGAGTG 436

QY 432 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 451

DB 437 GTTAGTCAAGAGCCTGTTTGTTCGGACACCATCATGTAACATATCAAGTATGGACGA 496

QY 452 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 471

DB 497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAATCGGTATGATTTT 556

QY 472 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 491  
DB 557 ATCATGAGAGTTTCTTAATAAATTTAATACATTTGGTAGGGAAAAAGAGCTCAATGAGT 616

QY 492 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 508

DB 617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTGCTGCTTAGTTGGAACCC 668

## RESULT 2

BM926413

LOCUS

DEFINITION

BM926413

VERSION

BM926413.1

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM926413 1033 bp mRNA linear EST 12-MAR-2002  
AGENCOURT 6600788 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5764845  
5', mRNA sequence.

ACCESSION BM926413

VERSION BM926413.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1033)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12818 row: o column: 22  
High quality sequence stop: 3  
High quality sequence stop: 686.

## FEATURES

source

1..1033

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5764845"

/clone\_lib="NIH MGC\_114"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: EcoRV (destroyed); RNA source anonymous pool of 6

male brains, age range 23-27 yo. Library is oligo-dT

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.5 kb, insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 019. Note:

this is a NIH\_MGC Library."

BASE COUNT 289 a 212 c 218 g 312 t

ORIGIN

Alignment Scores:

Pred. No.: 3,46e-120 Length: 1033

Score: 1068.50 Matches: 229

Percent Similarity: 93.17% Conservative: 3

Best Local Similarity: 91.97% Mismatches: 17

Query Match: 39.01% Indels: 4

DB: 14 Gaps: 0

US-09-873-409-7 (1-541) x BM926413 (1-1033)

QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaValProHisPheGlu 312

DB 180 GTTTTCTTAGTGTATCCATAGCAGTATTGATGGAGCAGAGTCCCTCACTTTGAA 239

QY 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332

DB 240 ACCTTCGCAATAGCCCGAGGAGTGGCTTTTCATATTTTCCAGGTTCATTGATAGAAACCC 299

```

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
DB 300 AGTATAGATAAATCTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTGGAA 359
QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeu 372
DB 360 TTTAAATAATGTTCTTCAATATATCCATCAGACCATCTATCAAGATTCCTGAAGGTCG 419
QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
DB 420 AATCTCAGAAATTAAGTCTCGAGAGACAGTCGCTTGGTGGTCTCAATGCGAGTGGGAAG 479
QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheLeuMetVal 412
DB 480 AGTACGGTAGTCAGCTTCTCGAGAGGTTATATGATCCGAGATGGCTTTATCATGGTG 539
QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
DB 540 GATGAGATGACATCAGACCTTTAAATGTGCGCATATTCGAGACCATATTCGAGTGGT 599
QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
DB 600 AGTCAAGAGCCTGTTTGTTCGGGACCATCAGTAACAATATCAAGTATGACGAGAT 659
QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
DB 660 GATGTGACTGATCAAGAGATGGAGAGACGAGCAAGGAAGCAAAATCGGTATGATTTATC 719
QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSerGly 492
DB 720 ATGGAGTTTCTTAATAA- TTTAATACATTTGGTAGGGGAANAAGGAGCTCCAAATGAGTGA 778
QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
DB 779 AGCCAGAACCG- AGGATCCCATTTGCTCTGCTTGTAGTTTGGAAACCCAGGAT- CTGAAT 836
QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 532
DB 837 TTAGATGAAGCTACGTCCTCCCTGGATTGAGAAACCAAGTCAACTGTGTCA- GCTGGCCTG 894
QY 533 GluLysAspThrProArgTyrSerPhe 541
DB 895 GAGAAGGAAACCCCGGNAATTCCTTT 921

RESULT 3
BM016204 603642859F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
LOCUS mRNA sequence.
DEFINITION BM016204.1 GI:16530558
ACCESSION EST.
VERSION BM016204.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 760)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DNP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLNL2068 row: e column: 16
High quality sequence stop: 740.
Location/Qualifiers

```

```

source 1..760
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5418615"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 229 a 150 c 187 g 193 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1,52e-115 Length: 760
Score: 1029.00 Matches: 209
Percent Similarity: 98.12% Conservative: 0
Best Local Similarity: 98.12% Mismatches: 2
Query Match: 37.57% Indels: 2
DB: 13 Gaps: 0

US-09-873-409-7 (1-541) x BM016204 (1-760)

QY 330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
DB 3 AAGAAACCCAGTATAGATAAATCTTCCACAGCTGGATATAAACCTGAATCCATAGAGGA 62
QY 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
DB 63 ACTGTGGAAATTTAAATAATGTTCTTTCAATATATCCATCAAGACCATCTATCAAGATTCG 122
QY 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
DB 123 AAAGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCTTGGTGGTCTCAATGGC 182
QY 390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
DB 183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTCGCAGAGGTTATATGATCCGAGTATGCTTT 242
QY 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
DB 243 ATCATGGTGGATGAGATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATN 302
QY 430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
DB 303 GGAGTGGTTAGTCAAGAGCTGTTTGTTCGGGACCATCATCAGTAACATATCAAGTAT 362
QY 450 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 469
DB 363 GGACCGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGGAAGCAAAATGCGTAT 422
QY 470 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGln 489
DB 423 GATTTTATCATCGAGTTCCTAATAAATTTAATACATTTGGTAGGGGAAAGAGAGCTCAA 482
QY 490 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 509
DB 483 ATGAGTGGAGGAGCAAAACAGAGGATCGCAATTTGGCTCGTCTAGTTTCGAAACCCCAA 542
QY 509 sileLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 529
DB 543 GATTTCTGATTTTCAGATGAGGCTAGCTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTT 602
QY 529 InAlaAlaLeuGluLysAspThrProArgTyrSer 540
DB 603 AAGCTGCATCTGGAGAAGGATACCCCGAGGTATTCC 637

RESULT 4
AL040762 AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS DKE2p434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION

```

DKFZp434C1815 5', mRNA sequence.  
 AL040762.1 GI:5409708  
 EST.  
 HOMO  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 545)  
 Bloecher, H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
 Wiemann, S.  
 EST (Bloecher, et al.)  
 Unpublished (1999)  
 Contact: Bloecher H  
 MIPS  
 Am Klopferpitz 18a D-82152 Martinsried, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
 Braunschweig/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.  
 si sequence also available.  
 This clone (DKFZp434C1815) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source  
 1..545  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp434C1815"  
 /clone\_lib="434 (synonym: htes3)"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /notes="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"  
 BASE COUNT 171 a 89 c 138 g 147 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4e-102 Length: 545  
 Score: 918.00 Matches: 181  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 33.52% Indels: 0  
 DB: 9 Gaps: 0

US-09-873-409-7 (1-541) x AL040762 (1-545)

QY 333 SerIleAsnValSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352  
 |||||  
 Db 3 AGTATAGATTAATCTTCCACAGCTGGATATTAACCTGATCCATAGAGACTGTGGAA 62  
 |||||  
 QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372  
 |||||  
 Db 63 TTTTAAATGTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCGTGAAGGTCGTG 122  
 |||||  
 QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392  
 |||||  
 Db 123 AATCTCAGAAATTAAGTCTGGAGACACAGTCGCTTGGTCGCTCAATGGCAGTGGGAAG 182  
 |||||  
 QY 393 SerThrValValGlnLeuGlnArgLeuTyrAspProAspGlyPheIleMetVal 412  
 |||||  
 Db 183 AGTACGGTAGTCCAGCTCTTCGACAGGTTATATGATCCGATGATGCTTTATCATGCTG 242  
 |||||  
 QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432  
 |||||  
 Db 243 GATGAGAATGACATCAGAGCTTTTAAATGTGGCGCATTTATCGAGACCATATTGGAGTGGTT 302  
 |||||  
 QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnIleLysTyrGlyArgAsp 452  
 |||||  
 Db 303 AGTCAAGAGCCTGTTTGTGTCGGACCAACCATCATGTAACAAATATCAAGTATGGACGAGAT 362  
 |||||

QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472  
 |||||  
 Db 363 GATGTGACTGATCAGAGATGGAGAGCAGCAAGCAAGCAAAATCGTATGATTTTATC 422  
 |||||  
 QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492  
 |||||  
 Db 423 ATGGAGTTTCTTAATAAATTTAATACATTGGTAGGGGAAAAGAGGCTCAATGAGTGGGA 482  
 |||||  
 QY 493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512  
 |||||  
 Db 483 GGGCAGAAACAGAGGATCGCAATTGCTGCTTGTAGTTGGAACCCCAAGATTCGATT 542  
 |||||  
 QY 513 Leu 513  
 |||||  
 Db 543 TTA 545  
 |||||  
 RESULT 5  
 BQ882401  
 LOCUS  
 DEFINITION BQ882401 929 bp mRNA linear EST 16-AUG-2002  
 AGENCOURT 8627902 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6291782  
 5', mRNA sequence.  
 ACCESSION BQ882401 GI:22274409  
 VERSION BQ882401.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 929)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L16M2492 row: c column: 15  
 High quality sequence stop: 677.  
 FEATURES  
 source  
 1..929  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6291782"  
 /clone\_lib="NIH\_MGC\_43"  
 /tissue\_type="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally-  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library. |"  
 BASE COUNT 281 a 176 c 237 g 234 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.7e-100 Length: 929  
 Score: 903.00 Matches: 188  
 Percent Similarity: 73.02% Conservatives: 61  
 Best Local Similarity: 55.13% Mismatches: 58  
 Query Match: 32.97% Indels: 34  
 DB: 14 Gaps: 3  
 US-09-873-409-7 (1-541) x BQ882401 (1-929)  
 QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThr-ValII 195

Db 1 CGGTATGCAAAAGCTGGACAGTAGCTGAAGAGGCTTGGCAGCAATTAGAAGCTGTGAT 60  
Qy 195 eAlaPheArgAlaGlnGlnGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 215  
Db 61 TCGATTGGAGACAAAGAGAACTTGAA----- 91  
Qy 215 rAlaTrpPheTyPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyTh 235  
Db 92 -----AGGTACAA 99  
Qy 235 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 255  
Db 100 CAAAATTTAGAGAGCTAAAGAAATTCGGATAAAGAAAGCTATTACAGCCAATATTTC 159  
Qy 255 rLeuGlyAlaValTyPhePheMetAsnGlyThrTyGlyLeuAlaPheTyrTyGlyTh 275  
Db 160 TATAGTGTCTTCT 219  
Qy 275 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 295  
Db 220 CACCTTGGTCTCTCAGGGGAA-----TATTCTATTGGACAAGTACTCACTGTAATCTT 273  
Qy 295 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 315  
Db 274 TTCTGTATTAATTGGGGCTTTTAGTGTGGACAGGCATCTCCAAGCAATTGAAGCATTTGC 333  
Qy 315 alIeAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 335  
Db 334 AAATGCAGAGGAGCGCTTATGAATCTTCAGATAATGTATGATAATGAAGCAAGTATGCA 393  
Qy 335 pAsnPheSerThrAlaGlyTyTyLysProGluSerIleGlyThrValGluPheLysAs 355  
Db 394 CAGCTATTGGAAGAGTGGGCACAAACAGATAATTAAGGGAAATTTGGAATTCAGAAA 453  
Qy 355 nValSerPheAsnTyProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuAr 375  
Db 454 TGTTCACCTTCAGTTACCCATCTCGAAAGAAAGTTAAGATCTTGAAGGGTCTGAACCTGAA 513  
Qy 375 rIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 395  
Db 514 GTGTCAGATGGGCACAGCGTGGCCCTGGTTTGGAAACAGCGTGGTGTGGGAAGAGCAAC 573  
Qy 395 lValGlnLeuLeuGlnArgLeuTyAspProAspAspGlyPheIleMetValAspGluAs 415  
Db 574 AGTCACAGCTGATGCAGAGGCTCTATGACCCACAGAGGGGATGGTCAGTGTTCATGGACA 633  
Qy 415 nAspIleArgAlaLeuAsnValArgHisTyArgAspHisIleGlyValValSerGlnGl 435  
Db 634 GGATATTAGACCAATAATGTAAGGTTTCTACCGGAAATCATTTGGTGTGTGAGTCAGGA 693  
Qy 435 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyGlyArgAspAspValTh 455  
Db 694 ACCTGTATTGTTTGGCCACACAGATGCTGAAACATCTCCATGCGCCGTGAAATGTCTAC 753  
Qy 455 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 475  
Db 754 CATGGATGAGATTGAGAAAGCTCTCCAGGAGCCATGCCCTATGATCTTATCATGAAC 813  
Qy 475 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly---G 494  
Db 814 TGCCTCATAAAATTGACACCCCTGGTTGGANAAAGAGGGGCCCAATTGAGTGTGTGGGGCAC 873  
Qy 494 lNlysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 511  
Db 874 AAACCAAGAGATCTCCCTTGGCAGCTGGGCCCTGGTTCCCAACCCCAAGATCCTC 928

RESULT 6  
LOCUS BM471690 948 bp mRNA linear EST 05-FEB-2002  
DEFINITION AGENCOURT\_6465349 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5539117  
5', mRNA sequence.  
ACCESSION BM471690

VERSION BM471690.1 GI:18520732  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 948)  
NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC/DCFD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12233 row: b column: 14  
High quality sequence stop: 569.  
FEATURES  
Location/Qualifiers  
1..948  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5539117"  
/clone\_lib="NIH\_MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: PCMV-SPORT6; Site: 1: Not I;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

BASE COUNT 285 a 187 c 194 g 282 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.64e-94 Length: 948  
Score: 856.50 Matches: 186  
Percent Similarity: 89.67% Conservative: 5  
Best Local Similarity: 87.32% Mismatches: 14  
Query Match: 31.27% Indels: 8  
DB: 13 Gaps: 1

US-09-873-409-7 (1-541) x BM471690 (1-948)

Qy 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312  
Db 205 GTTTCTTTAGTGAATCATCAGCAGTATTGCAATGGAGCAGCAGTCCCTCACTTTGAA 264  
Qy 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332  
Db 265 ACCTTCGCAATACCCGAGGAGCTGCTTTTCATATTTTCCAGGTATTGATAAGAAACCC 324  
Qy 333 SerIleAspAsnPheSerThrAlaGlyTyTyLysProGluSerIleGluGlyThrValGlu 352  
Db 325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAAATCCATAGAAGAACTGTGGAA 384  
Qy 353 PheLysAsnValSerPheAsnTyProSerArgProSerIleLysIleLeuLysGlyLeu 372  
Db 385 TTTTAAAAATTTCTTTTCAATATATCCATCAAGACCATCTATCAAGATTCGAAAGGTCTG 444  
Qy 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392  
Db 445 AATCTCAGATTAAGTCTGGAGAGACAGTCGCTTCGGTCTCAATGGCAGTGGGAAG 504  
Qy 393 SerThrValValGlnLeuLeuGlnArgLeuTyArgProAspAspGlyPheIleMetVal 412  
Db 505 AGTACGGTAGTCAGCTTCTGCAGAGGTTATATGATCCGATCCGATGATGGCTTATCATGGTG 564  
Qy 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAspHisIle-GlyValVa 432  
Db 565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTTGGAGTGGT 624

```

QY 432 iSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 451
|||||
Db 625 TAGTCAAGACCTCGTTTTCGGGACACCATCAGTAACATATCAAGTATGGAGC 684

QY 451 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 470
|||||
Db 685 AGATGATGGACTGATCAACAAAATGAAAAGCAACAGGGAAACCAAGGCGAATGAA 744

QY 471 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal--GlyGluLysGlyAlaGl 489
|||||
Db 745 TTTATCATGGAATTTCCCAAAAAATTTATATCATTTGGAAAGGGGAAAAAGACCTTCA 804

QY 489 nMetSerGlyGlyGlnLysGlnArg 497
|||||
Db 805 AGGAATGGAAGGCCCAAAACCCCAAG 829

RESULT 7
LOCUS BF584668 894 bp mRNA linear EST 12-DEC-2000
DEFINITION 602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
mRNA sequence.
ACCESSION BF584668
VERSION BF584668.1 GI:11658386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 894)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mai.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9798 row: 1 column: 02
High quality sequence stop: 651.
FEATURES
source
1..894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4218385"
/lab_host="NCI CGAP Co24"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 263 a 203 c 242 g 185 t
ORIGIN
1
Alignment Scores:
Pred. No.: 4,19e-81 Length: 894
Score: 750.00 Matches: 156
Percent Similarity: 84.94% Conservative: 47
Best Local Similarity: 65.27% Mismatches: 36
Query Match: 27.38% Indels: 4
DB: 12 Gaps: 0

US-09-873-409-7 (1-541) x BF584668 (1-894)

QY 293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
|||||
Db 5 GTCTTCTTTTCCTGTTAAT-GGAGCATTCAGTGTGGACAGGCATCTCCAAATATTGAA 63

```

```

QY 313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysPro 332
|||||
Db 64 GCCTTCGCCAATCCACGAGGACGCTTATGAAGCTTTCAAAATATTGATAATATGATCC 123

QY 333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
|||||
Db 124 AGTATAGACAGCTTCTCAAAAGAGTGGGCAACACCAACATACAGGAATCTGGAA 183

QY 353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
|||||
Db 184 TTTAAGATATTCACTTCAGTTACCATCTCGAAAAGAAGTTCAGATCTTGAAGGCGCTC 243

QY 373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
|||||
Db 244 AATCTGAAGGTGAAGAGCGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGGGAAA 303

QY 393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
|||||
Db 304 AGCACAACTGTCCAGCTGATCAAAAGGCTCTACGACCCCTAGATGGCATGGTCAGTAIC 363

QY 413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
|||||
Db 364 GACGACAGACATCAGAACCATCAATGTGAGGTATCTCGGGGAGATCATTTGTTGGTG 423

QY 433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
|||||
Db 424 AGTCAGAAACCTGTGCTGTTGTCACACGATCGCCGAGAACATTCGCTATGCCGAGAA 483

QY 453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
|||||
Db 484 GATGTCCACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCCAATGCCATGACTTCATC 543

QY 473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
|||||
Db 544 ATGAACCTGCCCAACCAATTGACACCTGTTGTTGGAGAGGGCGCAGCTGAGTGGG 603

QY 493 GlyGlnLysGlnArgIleAlaAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
|||||
Db 604 GGACAGAAACAGAGAATCGC-ATTGCCGGGCCCTGGTCCCAATCCCAAGTGCCTTC- 661

QY 513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 531
|||||
Db 662 GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCA 717

RESULT 8
LOCUS AK020318 872 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
(P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION AK020318.1 GI:12860872
VERSION AK020318.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male epididymis cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:9230106F14.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

```









```
VERSION AL520322.1 GH:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 253 a 225 c 253 g 211 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7,19e-70 Length: 943
Score: 659.00 Matches: 131
Percent Similarity: 72.36% Conservative: 47
Best Local Similarity: 53.25% Mismatches: 66
Query Match: 24.06% Indels: 2
DB: Gaps: 1
US-09-873-409-7 (1-541) x AL520322 (1-943)
QY 291 LeuAlaValPheSerValIleHisSerTyrCysIleGlyAlaAlaValProHis 310
Db 2 CTGTTATTATTTTCAGCTGTGTCTTTGGTCCATCGCGTGGGCAAGTCAGTTCAATTT 61
QY 311 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 330
Db 62 GTCCTCGACTATGCCAAAGCAAAATATCAGCAGCCACATCATCATGATCATGTAAGAAA 121
QY 331 LysProSerIleAspAenPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThr 350
Db 122 ACCCTTTGATTGACACTACAGCAGCGAGGCTATGCGACACATTTGGAGGGAAT 181
QY 351 ValGluPheLysAenValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys 370
Db 182 GTCACATTTTGGTGAAGTTGTATTCAACTATCCACCCGACCGGACATCCCGAGTCTCAG 241
QY 371 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 390
Db 242 GGAAGTGGCTGGAGGTGAAGAGGGGCACACGCTGTGCTTGGTGGGCAGCAGTGGCTGT 301
QY 391 GlyLysSerThrValValGlnLeuGlnAlaArgLeuTyrAspProAspAspGlyPheIle 410
Db 302 GGGNAGAGCACATGGTCTCCAGCTCTCGAGCGGGTCTACGACCCCTTGGCGGGAAGTG 361
QY 411 MetValAspGluAenAspIleAlaArgAlaLeuAenValArgHisTyrArgAspHisIleGly 430
Db 362 CTGCTTGATGGCAAGAAATAAAGCGACTGAATGTTCAAGTGGCTCCGAGCACACCTGGCG 421
QY 431 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 450
Db 422 ATCGTGTCCAGAGGCCCATCTCTTTTGGTCTGAGCATTTGCTGAGAACATTTGCTATGGA 481
QY 451 ArgAspAsp-----ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 468
Db 482 GACAACAGCGCGGTGTGTCTCAGAGGAAGAGATCGTGGGCGCAGCAAGAGGCGCAACATA 541
QY 469 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 488
Db 542 CATGCTTTCATCGAGTCACTGCTTAATAATATAGCACATAAGTAGGAGACAAGGAAC 601
QY 489 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 508
Db 602 CAGCTCTCTGTGGTGGCCAGAAACAACGCAATTCATAGCTCGTCCCTTGTAGACAGCCT 661
QY 509 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
Db 662 CATATTTCGTTTGGATGAAGCCAGTCAGCTCTGATACAGAAAAGTGAAGAGTTGTC 721
QY 529 GlnAlaAlaLeuGluLys 534
Db 722 CAAGAAGCCCTGGACAA 739
RESULT 11
LOCUS AA243820
DEFINITION zrf67906.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:668506 5',
similar to SWIMDR1_MOUSE P067905 MULTIDRUG RESISTANCE PROTEIN 1 ;,
mRNA sequence.
ACCESSION AA243820
VERSION AA243820.1 GI:1874631
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lemmon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,
T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
TITLE CONTACT: Wilson RK
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1572 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amerham.
FEATURES
Location/Qualifiers
1..405
/organism="Homo sapiens"
/db_xref="GDS:5562473"
/db_xref="taxon:9606"
/clone="IMAGE:668506"
/clone_lib="Soares NHMPu S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/notes="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NHMPu, and fetal heart NBH19W) were mixed, and as circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
```

5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 115 a 81 c 85 g 124 t  
ORIGIN

Alignment Scores:  
Pred. No.: 7,44e-70 Length: 405  
Score: 654.00 Matches: 128  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.22% Mismatches: 0  
Query Match: 23.88% Indels: 0  
DB: 9 Gaps: 0

US-09-873-409-7 (1-541) x AA243820 (1-405)

Qy 291 LeuAlaValPheSerValIleHisSerTyrCysIleGlyAlaValProHis 310  
Db 18 CTCTCTGTTTCTTTAGTGAATCCATAGCAGTTATTCATATTCAGGTTATTGATAAG 77  
Qy 311 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 330  
Db 78 TTTGAAACCTTCGCAATAGCCGAGGAGTGCCTTTCATATTTCCAGGTTATTGATAAG 137  
Qy 331 LysProSerIleAspAsnPheSerThrAlaGlyTyrIlyProGluSerIleGluGlyThr 350  
Db 138 AAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTCAATCCATAGAGAACT 197  
Qy 351 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys 370  
Db 198 GTGGAATTTTAAATGTTCTTTCAATATTCATCAAGACCATATCAAGATTCAGAA 257  
Qy 371 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 390  
Db 258 GGTCTGAATCTCAGATTAAGTCTGGAGACAGTCCGCTTGGTCTCTCAATGGCAGT 317  
Qy 391 GlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIle 410  
Db 318 GGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATGATCCGATGATGCTTTATC 377  
Qy 411 MetValAspGluAsnAspIleArgAla 419  
Db 378 ATGGTGATGAGATGACATCAGAGCT 404

RESULT 12  
BJ075336  
LOCUS  
DEFINITION  
BJ075336 NIBB Mochii normalized Xenopus tailbud library  
ACCESSION  
BJ075336.1 GI:17520252  
VERSION  
EST.  
KEYWORDS  
SOURCE  
ORGANISM  
African clawed frog.

550 bp mRNA linear EST 11-DEC-2001  
BJ075336 NIBB Mochii normalized Xenopus tailbud library  
laevis cdna clone XL057d10 5', mRNA sequence.

Expressed genes in X. laevis embryo  
Unpublished (2001)  
Contact: Tadao Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1. .550  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"

REFERENCE  
AUTHORS  
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara  
,Y.

TITLE  
JOURNAL  
COMMENT

FEATURES  
source

/clone="XL057d10"  
/clone\_lib="NIBB Mochii normalized Xenopus tailbud  
library"

/tissue\_type="whole embryo"  
/dev\_stage="stage 25"

/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs  
were oligo-dT primed and directionally cloned. Staging  
according to Nieuwkoop and Faber. Library is subtracted  
and was constructed by N. Garrett and A.M. Zorn,  
(Wellcome/CRC Institute)."

BASE COUNT 162 a 109 c 144 g 135 t  
ORIGIN

Alignment Scores:

Pred. No.: 3.81e-69 Length: 550  
Score: 650.00 Matches: 121  
Percent Similarity: 85.79% Conservative: 36  
Best Local Similarity: 66.12% Mismatches: 26  
Query Match: 23.73% Indels: 0  
DB: 13 Gaps: 0

US-09-873-409-7 (1-541) x BJ075336 (1-550)

Qy 349 GlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIle 368  
Db 2 GCGCATATAGAAATTCAAAGGCATTTCTTTGCTACCCATGCCGCTCTGCTGTGCAAAAT 61

Qy 369 LeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsn 388  
Db 62 CTCTCTGGACTGAACTGAAAGTGGAGCTGTAAACCATCGCTTTGGTGGTATGAGT 121

Qy 389 GlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGly 408  
Db 122 GGTGTGTGTAAGAAAGCACTACCATACAGTTGTGTCAGAGGTTCTATCATCATCAGATGGA 181

Qy 409 PheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHis 428  
Db 182 GAGGTTACTTTGGATGGAGCAGCATATCCGTTCACTAAATGTAATAATGGTTGAGGGAAT 241

Qy 429 IleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLys 448  
Db 242 ATTGGGTAGTAGTGCAGGAGCCAGTCTCTTTGGCACAACCATTCGAGAAATATATCCGC 301

Qy 449 TyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 468  
Db 302 TACGGAGGGGATCAGTAACAGAACAAATTTGAACAAGCGGCAAGCAATGCC 361

Qy 469 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAla 488  
Db 362 TTTGATTTTATATCCAGGCTTCTCTGATAAATTTAAGACGATGGTGGAGAAAGAGTGGC 421

Qy 489 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 508  
Db 422 CAGTTGAGTGGGGTTCAGAGCAGCGATAGCCATTCAGCCGCTCTGTACGAACCTT 481

Qy 509 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528  
Db 482 AAGATTTCTACTTCTTGATGAGCCACATCAGCTCTCGATACACAAAGTGAAGCTATTGTA 541

Qy 529 GlnAlaAla 531  
Db 542 CAAGCTGCC 550

RESULT 13  
BG298756  
LOCUS

DEFINITION  
BG298756 602396681F1 NIBB Mochii normalized Xenopus tailbud library  
ACCESSION  
BG298756

KEYWORDS  
SOURCE  
ORGANISM  
house mouse.

803 bp mRNA linear EST 21-FEB-2001  
BG298756 602396681F1 NIBB Mochii normalized Xenopus tailbud library  
mus musculus cdna clone IMAGE:451157 5',  
mRNA sequence.

ACCESSION  
BG298756

KEYWORDS  
SOURCE  
ORGANISM  
house mouse.

803 bp mRNA linear EST 21-FEB-2001  
BG298756 602396681F1 NIBB Mochii normalized Xenopus tailbud library  
mus musculus cdna clone IMAGE:451157 5',  
mRNA sequence.

ACCESSION  
BG298756

KEYWORDS  
SOURCE  
ORGANISM  
house mouse.

803 bp mRNA linear EST 21-FEB-2001  
BG298756 602396681F1 NIBB Mochii normalized Xenopus tailbud library  
mus musculus cdna clone IMAGE:451157 5',  
mRNA sequence.

ACCESSION  
BG298756

KEYWORDS  
SOURCE  
ORGANISM  
house mouse.

803 bp mRNA linear EST 21-FEB-2001  
BG298756 602396681F1 NIBB Mochii normalized Xenopus tailbud library  
mus musculus cdna clone IMAGE:451157 5',  
mRNA sequence.

ACCESSION  
BG298756

KEYWORDS  
SOURCE  
ORGANISM  
house mouse.

803 bp mRNA linear EST 21-FEB-2001  
BG298756 602396681F1 NIBB Mochii normalized Xenopus tailbud library  
mus musculus cdna clone IMAGE:451157 5',  
mRNA sequence.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 803)  
 NIH-MGC <http://mgi.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Plate: LLAM10394 row: b column: 22  
 High quality sequence stop: 698.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

1. .803  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone\_lib="IMAGE:4511157"  
 /tissue\_type="retina"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."  
 BASE COUNT 240 a 162 c 205 g 196 t

ORIGIN

Alignment Scores:  
 Pred. No.: 1,25e-68 Length: 803  
 Score: 648.00 Matches: 149  
 Percent Similarity: 68.99% Conservative: 49  
 Best Local Similarity: 51.92% Mismatches: 55  
 Query Match: 23.66% Indels: 37  
 DB: 12 Gaps: 3

US-09-873-409-7 (1-541) x BG298756 (1-803)

Qy 168 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluVal 187  
 Db 15 TCATTACTGATAGGAACCTCCATGCTTATGCAAAAGCTGAGCAGTGTCTGAGAAGTC 74  
 Qy 188 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 207  
 Db 75 TTACGACCATCAGAACTGTGATGTCGTGGA-CGACAAAAGGAAGCACTTGAA----- 127  
 Qy 208 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 227  
 Db 127 ----- 127  
 Qy 228 ValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 247  
 Db 128 -----AGTCATATACAACTTGAAGAAGCACTTGAAGAAGCTGAGGATTAAG 172  
 Qy 248 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 267  
 Db 173 AAGCTATCAGCGCCAACTCTCATGGTGCAGCTTTCTCTTATCATGATCATAT 232  
 Qy 268 GlyLeuAlaPheTrpTyrGlyThrSerLeuLeuLeuAsnGlyGluProGlyTyrThrIle 287  
 Db 233 GCTCTGGCAATTCGTATGGAGCTTCCTTGTGTCATCTCCAAAGAA-----TACTCTAT- 285  
 Qy 288 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 307  
 Db 286 GGACAAGTGCCTACTGCTCTCTTTCCGTGTAAATGAGCAATTCATGTTGGACAGGCA 345  
 Qy 308 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 327  
 Db 308 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 327

Db 346 TCTCCAATATTGAAGCCTTCGCCATGCAGGAGCAGCTTATGAAGTCTTCAAAATA 405  
 Qy 328 IleAspLysIleValProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 347  
 Db 406 ATTGATAATAAGCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACAGACAATA 465  
 Qy 348 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 367  
 Db 466 CRAAGAAATCTGGAATTTAAGATATTCACTTCACTTCCATCTCGAAAAGAGTTTCAG 525  
 Qy 368 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 387  
 Db 526 ATCTTCAGAGCCCTCAATCTGAAGTGAAGAGCGGACAGCGTGGCCCTGTGGTGAAC 585  
 Qy 388 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 407  
 Db 586 AGTGGCTGTGGAAAAAGCACACTGTCCAGCTGATCCAAGGCTCTACGACCCTA-GAT 644  
 Qy 408 GlyPheIleMetValAspGluAsnAspIleArgAlaLeu-AsnValArgHis-TyrArg- 426  
 Db 645 GGCATGTCAGTATCCAGCGACGACATCAGAACCATCAATGTGAGGTATCTGGAGGG 704  
 Qy 427 AspHisIleGlyValValSerGln-GluProValLeuPheGlyThrThrIleSerAsnAs 446  
 Db 705 GATCAT--GGGGTGTGAGTCAAGGAACCTGGCTTTGTTGCCAACCCAGATCGGAGAA 761  
 Qy 446 nleLysTyrGly 450  
 Db 762 AATTCGTTATGGG 774

RESULT 14

LOCUS

BG587938 780 bp mRNA linear EST 11-APR-2001  
 EST489713 MHAM Medicago truncatula/Glomus versiforme mixed EST  
 library cDNA clone pMHAM-51L19 5' end, mRNA sequence.

ACCESSION

BG587938

VERSION

BG587938.1 GI:13603002

KEYWORDS

EST.

SOURCE

Medicago truncatula/Glomus versiforme mixed EST library.  
 Medicago truncatula/Glomus versiforme mixed EST library.  
 Eukaryota; mixed EST libraries.

REFERENCE

1 (bases 1 to 780)  
 Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.

AUTHORS

and Fraser,C.M.

ESTs from roots of Medicago truncatula after colonization with

TITLE

Glomus versiforme, 2001

JOURNAL

Unpublished (2001)

COMMENT

Contact: Harrison M.J.

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73401

Tel: 580-223-5810

Fax: 580-221-7380

Email: [mj.harrison@noble.org](mailto:mj.harrison@noble.org)

Noble EST name: N387011e TIGR sequence name: MTDDN70TK More  
 information is available at: <http://www.medicago.org>  
 Seq primer: SKnod (CTA gaa cta gta gat cc).

FEATURES

source

1. .780  
 /organism="Medicago truncatula/Glomus versiforme mixed EST  
 library"  
 /cultivar="Medicago truncatula genotype A17"  
 /db\_xref="taxon:119092"  
 /clone\_lib="MHAM-51L19"  
 /tissue\_type="roots"  
 /note="Roots colonized with Glomus versiforme"  
 /dev\_stage="Roots harvested at 10, 17, 22, 31 and 38 days  
 post-inoculation with Glomus versiforme. The library was  
 made from a mixture of RNA from each of these stages."  
 /lab\_host="E. coli strain XL0LR"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA from  
 roots harvested at 10, 17, 22, 31 and 38 days







GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 79.2267 Seconds

(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739

Sequence: 1 MILGTILASLVNGACILPLMLP.....SESKSAVQALEKDTFRYSF 541

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO/spool/US09873409/runat\_27032003\_115422\_19259/app\_query\_fasta\_1.7544  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 578 @runat\_27032003\_115422\_19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552.5	56.7	4233	3 US-09-120-513-1	Sequence 1, Appli
2	1552.5	56.7	4233	4 US-09-450-105-1	Sequence 1, Appli
3	1552.5	56.7	4669	2 US-08-752-447-1	Sequence 1, Appli
4	1552.5	56.7	4669	4 US-09-316-167-1	Sequence 1, Appli
5	1546.5	56.5	4646	1 US-08-181-471-2	Sequence 2, Appli
6	1546.5	56.5	6505	2 US-08-793-610-5	Sequence 5, Appli
7	1546.5	56.5	9318	2 US-08-793-610-6	Sequence 6, Appli
8	1545	56.4	4264	2 US-08-784-649A-1	Sequence 1, Appli
9	1545	56.4	4264	2 US-08-784-649A-5	Sequence 5, Appli
10	1544.5	56.4	4669	6 5206352-3	Patent No. 5206352
11	1525.5	55.7	4669	2 US-08-583-276-18	Sequence 18, Appli
12	960	35.0	4224	1 US-08-612-521-1	Sequence 1, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-120-513-1

; Sequence 1, Application US/09120513

; Patent No. 6025160

; GENERAL INFORMATION:

; APPLICANT: Brun, Kimberly

; APPLICANT: Chenery, Richard

; APPLICANT: Ellens, Harma

; APPLICANT: Field, John

; APPLICANT: Yue, Lin

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE

; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY:

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: Fast-Seq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/120,513

; FILING DATE: 22-JUL-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: King, William T

; REGISTRATION NUMBER: 30,954

; REFERENCE/DOCKET NUMBER: GPO008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5015

13	935.5	34.2	4047	2	US-08-612-734B-1	Sequence 1, Appli
14	916	33.4	4800	2	US-08-612-734B-3	Sequence 3, Appli
15	914	33.4	4002	2	US-08-996-545-1	Sequence 1, Appli
16	914	33.4	4002	2	US-08-996-545-3	Sequence 3, Appli
17	914	33.4	4002	4	US-09-328-320-1	Sequence 1, Appli
18	914	33.4	4002	4	US-09-328-320-3	Sequence 3, Appli
19	898	32.8	2726	1	US-08-461-823-1	Sequence 1, Appli
20	897.5	32.8	3924	1	US-08-395-246C-1	Sequence 1, Appli
21	849.5	31.0	6143	1	US-08-612-521-3	Sequence 3, Appli
22	785.5	28.7	3792	4	US-09-351-224E-10	Sequence 10, Appli
23	747.5	27.3	3924	2	US-08-996-644-3	Sequence 3, Appli
24	747.5	27.3	3924	3	US-09-352-552-3	Sequence 3, Appli
25	747.5	27.3	3927	2	US-08-996-644-1	Sequence 1, Appli
26	747.5	27.3	3927	3	US-09-352-552-1	Sequence 1, Appli
27	688.5	25.1	3999	4	US-09-351-224E-9	Sequence 9, Appli
28	646	23.6	2376	1	US-08-394-880B-1	Sequence 1, Appli
29	626.5	22.9	2061	4	US-09-061-764A-17	Sequence 17, Appli
30	626	22.9	3909	1	US-08-232-537-1	Sequence 1, Appli
31	621.5	22.7	1959	4	US-09-061-764A-4	Sequence 4, Appli
32	574.5	21.0	2244	4	US-09-061-764A-18	Sequence 18, Appli
33	571.5	20.9	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
34	529	19.3	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
35	519.5	19.0	7760	4	US-08-961-527-63	Sequence 63, Appli
36	501.5	18.3	5120	3	US-08-772-270A-6	Sequence 6, Appli
37	501.5	18.3	8370	2	US-08-488-706-1	Sequence 1, Appli
38	491	17.9	4411529	4	US-09-103-840A-1	Sequence 1, Appli
39	489	17.9	13188	4	US-08-961-527-70	Sequence 70, Appli
40	489	17.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
41	483	17.6	7721	3	US-08-772-270A-14	Sequence 14, Appli
42	482	17.6	28804	2	US-08-592-874-1	Sequence 1, Appli
43	482	17.6	28804	3	US-09-096-942-2	Sequence 2, Appli
44	482	17.6	28804	3	US-09-096-867-2	Sequence 2, Appli
45	460	16.8	6492	4	US-08-961-527-188	Sequence 188, App

; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4233 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
US-09-120-513-1

## Alignment Scores:

Pred. No.: 1.69e-172 Length: 4233  
Score: 1552.50 Matches: 308  
Percent Similarity: 72.20% Conservative: 92  
Best Local Similarity: 55.60% Mismatches: 105  
Query Match: 56.68% Indels: 49  
DB: 3 Gaps: 5

US-09-873-409-7 (1-541) x US-09-120-513-1 (1-4233)

Qy 1 MetileuGlylleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 170 ATGGCTCTGGAACTCTCGCTGCTATCATCCAGGAACCCCTGCTTCCCTCCTGATGCTG 229  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30  
Db 230 GTGTTCGATACATGACAGATAGTTTACCACAGCAGAGCCCGCATCTGCCGAGGCTT 289  
Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45  
Db 290 ACTAATCAAGTGAATCAACAGTACACAGACCGCTCAGCAGACAGCATCTGGAGGAGAC 349  
Qy 46 -----ThrLeuTyrTyrValGlylleGlyValAlaLeuLeuPheGlyTyr 61  
Db 350 ATGGCCATGTACGCTACTATTACACGGGCATTTGGTCCGCTGTGCTCATCTGCCCTAC 409  
Qy 62 ileGlnIleSerLeuTyrPilelleThrAlaAlaArgGlnThrLysArgIleArgGln 81  
Db 410 ATCCAGGTTTCACCTTGGTGGCTGGAGCTGGGAGACAAATACACAGATTTAGCAGAAG 469  
Qy 82 PhePheHisSerValLeuAlaGlnAspileGlyTrpPheAspSerCysAspileGlyGlu 101  
Db 470 TTTTTCATGCCATCATGAATCAGAGATAGCTGGTTTGAACGTGAATGACGCTGGGAG 529  
Qy 102 LeuAsnThrArgMetThr---AspileAspLysIleSerAspGlylleGlyAspLysIle 120  
Db 530 CTCAACACCCGGCTCACAGATGACGCTCTCCAAAATTAATGACGGAATTTGGTGACAACTT 589  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 590 GGAATGTTCTTTCAGTCATCAACAGACATTTTACGCCGGTTTATTAATAGATTTATAGT 649  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 650 GGTGGAGCTTAACCTTGTAAATTTGGCCGCTCAGCCCTCTTATTGGGTGCTCATCTGCC 709  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 710 ATGTGGGCAAGGTAAGTACTGACTTCAATTACTAATAGGAACCTCCAGGCTTATGCCAAAGCT 769  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 770 GGACAGTTTGGCGAAGAGTCTTAGCAGCCATCAGACTGTGATGGTTGGAGGACAA 829  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgfyrAlaTrpPheTyrPhe 220  
Db 830 AAGAAGGAACCTTGAA----- 844  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgfyrThrGlnAsnLeuLysAsp 240  
Db 845 -----AGGTACAATAAAATTTAGAGAA 868

Qy 241 AlaLysAspPheGlylleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 869 GCTAAAGAGTTGGCATAAAGAAAGCCATCAGCGCAACATTTCCATAGTATTGCTAC 928  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 929 CTGTTGGTCTATGCTGCTTATGCACTGGCATTTCTGTATGGGACCTCTTGGTCTCTCA 988  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 989 AATGAA-----TATTCTATTGGCAAGTCTTACCGTCTTCTCTCTATTATTATGGGG 1042  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1043 ACITTCAGTATTGGACATTTAGCCCCCAACATAGAGCCTTTGCAAAATGCAAGAGGGCA 1102  
Qy 321 AlapheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCCAAGCATTCACACGCTTCTCAACCAG 1162  
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1163 GGACAAACACACAGATATATGGGAATTTGGAATTTTAAATGTTTACTTCAACTAC 1222  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1223 CCATCAGAAAGTCAAGTTAAGATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGGCAG 1282  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1283 ACGGTACGCCCTGTTGGCAACAGTGGCTGTGGGAAAGCAACCTGTCTCCAGCTGCTGCAG 1342  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspileArgAlaLeu 420  
Db 1343 AGGCTCTACACCCCATAGAGGGCGAGTCACTATCAGCGACAGGACATCAGACCATC 1402  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1403 AATGTAGGTTATCTCGGGAATCAATTTGGGTGTGTGAGTCAGGAACCCGTGTTTGGC 1462  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1463 ACCACGATTTCCGAAACATTCGCTATGGCCGAGAAACCTCCCATGGATGAGATAGAG 1522  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1523 AAGCTGTCAAGGAAGCCATGCTATGACTTATATGAACCTGCCCCCAAAATTTTAC 1582  
Qy 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1583 ACCCTGTTGTCAGAGAGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1643 GCCCGGCGCTGTGCGCAACCCCAAGATCTTTTGTGTGATGAGGCCACGTCAGCCTTG 1702  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1703 GACACAGAAAGCAAGCCGTTGTTTACGGCCCTCTCTGGATAG 1744

## RESULT 2

US-09-450-105-1  
; Sequence 1, Application US/09450105  
; Patent No. 6169166  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Ellens  
; APPLICANT: John Anthony Feild  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF  
; FILE REFERENCE: GP-50008-D1

; CURRENT APPLICATION NUMBER: US/09/450,105

; CURRENT FILING DATE: 1999-11-29

; EARLIER APPLICATION NUMBER: 09/120,513

; EARLIER FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 4233

; TYPE: DNA

; ORGANISM: HOMO SAPIENS

US-09-450-105-1

#### Alignment Scores:

Pred. No.: 1.69e-172 Length: 4233  
Score: 1552.50 Matches: 308  
Percent Similarity: 72.20% Conservative: 92  
Best Local Similarity: 55.60% Mismatches: 105  
Query Match: 56.68% Indels: 49  
DB: 4 Gaps: 5

US-09-873-409-7 (1-541) x US-09-450-105-1 (1-4233)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 170 ATGGCTCTGGGAACCTCGCTGCTATCATCCACGGAACCTGCTCCCTCTGTGCTG 229  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30  
Db 230 GTGTCGGATACATGACATAGATTTTACCAAGCAGACAGCCGCTTCTGCCGAGCGTT 289  
Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPheArgLeu----- 45  
Db 290 ACTAATCAAAGTGAATCAACAGTACACAGCCGTCAGCAGCAGCAGTCTGGAGGAGGAC 349  
Qy 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaLeuLeuPheGlyTyr 61  
Db 350 ATGGCCATGTACGCCCTACTATTACAGCGGCAATGGTGGCGGTGTCTCATCGTTCCTAC 409  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 410 ATCCAGGTTTCATTGGTCCCTGGCAGCTGGGAGACAAATACACAAGATTAGCCAGAG 469  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 470 TTTTTCATGCCATCATGAATCAGGAGATAGGCTGTTGACCTGAATGACGCTGGGAG 529  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 530 CTCACACCCGGCTCAGATGATGATGCTCCAAAATTAATGACGGAATTTGTTGACAACTT 589  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 590 GGAATGTCTTTTCAGTCCATACACATTTTACCGCGTGTATATAGATTTATAGT 649  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 650 GGTGTGAAGTACCTACCTGTAAATTTGGCGTTCAGCCCTCTTATTTGGTGTTCATCTGCC 709  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180  
Db 710 ATGTGGGCAAGGTACTGACTTCTATTACTAATAAGGAACCTCCAGGCTTATGCGAAGCT 769  
Qy 181 GlyValAlaValGlnGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 770 GGAGCAGATTGCCAAGAAGTCTTAGCAGCCATCAGAACTGTGATGCTGTTGGAGGACAA 829  
Qy 201 GluLysGlnLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 830 AAGAGGAATCTGAA----- 844  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 845 -----AGGTACAATAAAATTTAGAGAA 868

Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 869 GCTAAAGAGATTGGCATAAAGAACCCATCAGGGCCAAACATTTCCATAGGTATGCTTAC 928  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280  
Db 929 CTGTTGGTCTATGCGTCTATGCACTGGCATCTCTGGTATGGACCTCCTTGGTCTCTCA 988  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 989 AATGAA-----TATTCTATTGGACAAGTGTCTTACCGTCTTCTCTCTATTTATTGGGG 1042  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1043 ACTTTCAGTATTGGACATTTAGCCCAACATAGAGCCCTTTGCAAAATCAAGAGGGGCA 1102  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCCAAGCATTTGACAGCTTCTCAACAAG 1162  
Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1163 GCACACAAACACAGACAGTATATGGAATTTGGAATTTAAATAATGTTACTTCAACTAC 1222  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1223 CCATCAGGAAGTGAAGTAAAGATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGCGAG 1282  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
Db 1283 ACGGTAGCCCTGTTGGCAACAGTGGCTGTGGAAAGCACAACTGTCCAGCTGTGCGAG 1342  
Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
Db 1343 AGGCTCTACACACCCATAGAGGGCGAGGTCTAGTATCGACGGACAGACATCAGGACCATC 1402  
Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
Db 1403 AATGTGAGGTATCTCGGGAATCATTTGGGGTGGTGGTTCAGGAACCCGCTGTGTTGCC 1462  
Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
Db 1463 ACCACGATTTGCCGAAACATCTCGTATGCGCCGAGAAAGCTCACCATTGATGATAGTAG 1522  
Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
Db 1523 AAGCTGTCAAGAGCCCAATGCTATGCTATGCTCATCATGAACTGCCCCACAAATTTAAC 1582  
Qy 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
Db 1583 ACCCTGTTGGTGGAGAGGGCGGCGAGTGTGGGGGACAGAAACAGAGGATCGCCATT 1642  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 1643 GCCCGGCGCTGCTGCGCAACCCCAAGATCCTTTTGTGGATGAGGCGCACGTCAGGCTTG 1702  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534  
Db 1703 GACACAGAAAGCGAAGCCGCTGTTTACGGCGCTCTCGGATAAG 1744

#### RESULT 3

US-08-752-447-1

; Sequence 1, Application US/08752447

; Patent No. 5994088

; GENERAL INFORMATION:

; APPLICANT: Mechneter, Eugene

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods and Reagents for Preparing and

; TITLE OF INVENTION: Using Immunological Agents Specific for p-glycoprotein

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.

STREET: 300 South Wacker Drive, Seventh Floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/752,447  
 FILING DATE: 15-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 5994088nan, Kevin E  
 REGISTRATION NUMBER: 35,303  
 REFERENCE/DOCKET NUMBER: 95,1121  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-913-0001  
 TELEFAX: 312-913-9808  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4669 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: 5'UTR  
 LOCATION: 1..424  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 425..4264  
 FEATURE:  
 NAME/KEY: 3'UTR  
 LOCATION: 4265..4669  
 US-08-752-447-1

Alignment Scores:  
 Pred. No.: 1.99e-172 Length: 4669  
 Score: 1552.50 Matches: 305  
 Percent Similarity: 73.29% Conservative: 101  
 Best Local Similarity: 55.05% Mismatches: 99  
 Query Match: 56.68% Indels: 49  
 DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-752-447-1 (1-4669)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 575 ATGGTGGGAACTTTGGCTGCCATCATCCATGGGCTGCACCTCTCTCATGATGCTG 634  
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGAGGAAATTTAGAGATCTCATGTCA 694  
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 695 AACATCAATAAGAGATGATCAATGATACAGGGTTCTTTCATGAATCTGGAGGAAGAC 754  
 Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 755 ATGACCGATGATGCTATTATTCAGTGAATTTGGCTGGGGTGGCTGGTGGCTGCTTAC 814  
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgGln 81  
 Db 815 ATTGAGGTTTCATTGGTGGCTGGCAGCTGGAAGACAAATACACAAAATTTAGAAAACAG 874  
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
 Db 875 TTTTTCATGCTATTATCGCAGGAGATAGCTGGTTTGTGATGTCACGATGTTGGGAG 934

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 935 CTTTAAACCCGACTTACAGATGATGCTCCAGATTAATGAAGGAATTTGGTGAACAAAT 994  
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054  
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 1055 GGTGGAAGCTAACCTTGGATTTGGCCATCAGTCTCTGTTCTGGACTGTGAGTGTCT 1114  
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaIleLysAla 180  
 Db 1115 GTCTGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174  
 Qy 181 GlyAlaValAlaGlnValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 1175 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATTGCAATTGGAGGCAA 1234  
 Qy 201 GluLysGlnLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
 Db 1235 AAGAAAGAACTTGAA----- 1249  
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
 Db 1250 -----AGGTACAACAAAAATTTAGAGAA 1273  
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
 Db 1274 GCTAAAAGAAATTTGGATAAAGAAAGCTATTACAGCCAAATTTCTATAGTGTGCTTTC 1333  
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
 Db 1334 CTGCTGATCTATGCATCTTAAGTCTGGCCCTCTGGTATGGACCACTTGGTCTCTCA 1393  
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 Db 1394 GGGAA-----TATCTATTGGCAAGTACTCACTGTTCTTTCTGTTATTAAATGGG 1447  
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
 Db 1448 GCTTTTAGTGTGGCAGGCACTCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGAGCA 1507  
 Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
 Db 1508 GCTTATGAATCTTCAAGATAATTGATAAGCAAGTATTGACAGCTATTGCAAGAGT 1567  
 Qy 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
 Db 1568 GGGCAACAAACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627  
 Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
 Db 1628 CCATCTCGAAAAGAAAGTTAAGATCTTGAAGGGCTTGAACCTGAAGGTGCGAGAGTGGCAG 1687  
 Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
 Db 1688 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGAGAGCAACACAGTCCAGCTGATGCGAG 1747  
 Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
 Db 1748 AGGCTCTATGACCCACAGAGGGGATGGTTCAGTGTTCATGGACAGGATATTAGGACCATA 1807  
 Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
 Db 1808 AATGTAAGGTTTCTACGGGAATCATTTGGTGTGGTGTAGTCAGGAACCTGATTGTTGGCC 1867  
 Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460  
 Db 1868 ACCACGATATCGTGAACATTTGGTATGGCCGTGAAATGTCACCATGGATGAGATTGAG 1927  
 Qy 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480

```
Db 1928 AAAGCTGTCAAGGAAGCCCAATGCCTATGATCTTATCATGAAGCTGCCTCATAAATTGAC 1987
Qy 481 ThrLeuValGlyGlyGlyAlaGlnMetSerGlyGlyGlnGlyGlnAlaIle 500
Db 1988 ACCCTGGTGGAGAGAGAGGGCCAGTGGTGGTGGGCGAGAGAGAGATCGCCATT 2047
Qy 501 AlaArgAlaLeuValArgAnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db 2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTGAGCCTTG 2107
Qy 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db 2108 GACACAGAAAGCAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149

RESULT 4
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechneter, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; NUMBER OF INVENTION: Using Immunoligal Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6365357man, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.: 1,99e-172 Length: 4669
Score: 1552.50 Matches: 305
Percent Similarity: 73.29% Conservative: 101
```

```
Best Local Similarity: 55.05% Mismatches: 99
Query Match: 56.68% Indels: 49
DB: 4 Gaps: 6
US-09-873-409-7 (1-541) x US-09-316-167-1 (1-4669)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGTGGGAACCTTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTGGAGAAATGACATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGTCA 694
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCAGCTAATAGAGTGCATATCAATGATACAGGGTCTTCTCATGAATCTGGAGGAGAC 754
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAGGTATGCCTATTATTACAGTGGAAATTTGGTCTGGGGTGTCTGGTGTCTTAC 814
Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGTTTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGTATGTGCACCATGTTGGGGAG 934
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCCAAGATTAAATGAAGGAATTTGGTGACAAAT 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTCAGGTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACTATCTTCACTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTCGATTTGGAGGACAA 1234
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTyrPheTyrPhe 220
Db 1235 AAGAAAGAACTTGAA----- 1249
Qy 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db 1250 -----AGGTCAACAAATAATTTAGAGAA 1273
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATTTTCTATAGGTGCTGCTTTC 1333
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
Db 1334 CTGCTGATCTATGATCTTATGCTGGGGCTTCTGGTATGGGACCACTTGGTCTCTCA 1393
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db 1394 GGGAA-----TATTCTATTGGACAAGTACTCAGTGTATTCTTTCTGATTAAATGGG 1447
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db 1448 GCTTTTAGTGTGGACAGGCATCTCCAAAGCATTTGAAGCATTTTGCAAAATGCAAGAGGACGA 1507
```









MOLECULE TYPE: DNA  
US-08-793-610-6

## Alignment Scores:

3.25e-171 Length: 9318  
Score: 1546.50 Matches: 305  
Percent Similarity: 73.10% Conservative: 100  
Best Local Similarity: 55.05% Mismatches: 100  
Query Match: 56.46% Indels: 49  
DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-793-610-6 (1-9318)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAlaGlnGlyAlaCysLeuProLeuMetProLeu 20  
DB 1926 ATGGTGGTGGAACTTTGGTGCCTCATCCATGGGGTGGACTTCTCTCATGATGCTG 1985  
QY 21 ValLeuGlyGluMetSerLeu-----AsnLeuIleSer 31  
DB 1986 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCGAGAAATTTAGAAATCTGATGTC 2045  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 2046 AACATCACTAATAGAGATGATATCATATGATGATGATGATGATGATGATGATGATGATG 2105  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 2106 ATGACAGGATGCTTATATATACAGTGAATTTGGTGGGTGCTGCTGCTGCTTAC 2165  
QY 62 IleGlnIleSerLeuTrpIleLeuAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 2166 ATTCAGGTTTCTTTGGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 2226 TTTTTCATGCTATATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2285  
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 2286 CTTAACACCGACTTACAGATGATGCTCTAAGATTAATGAGATTTGCTGACAAATTT 2345  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 2346 GGAATGTTCTTTTCTCAGTCAATGCGAACATTTTCTCAGTGGTATTATGATGATGATGAT 2405  
QY 141 GlyTriPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 2406 GGTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTTGGACTGTGCTGCTGCT 2465  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 2466 GTCTGGCAAGATATCTATCTTATTACTGATTAAGAACTCTTAGCTATGCAAAAGCT 2525  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 2526 GGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAATCTGATGATGATGATGATGATGAT 2585  
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
DB 2586 AAGAAAGAACTTGAA----- 2600  
QY 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
DB 2601 -----AGGTACAAACAAAAATTTAGAGAA 2624  
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
DB 2625 GCTAAAGAAATTTGGATAAGAAAGCTTATTACAGCAATATTCTATAGTGTGCTGCTTC 2684  
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
DB 2685 CTGCTGATCTATGATCTTATGCTCTGGCTCTGCTGATGGGACCACTTGGTCTGCTCTCA 2744

QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
DB 2745 GGGGAA-----TATTTCTATGGCAAGTACTCCTGTTATTTCTCTGTTATTAATGGG 2798  
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
DB 2799 GCTTTTAGTGTGGACAGGATCTCAAGCATTTGAAGCATTTGCAATCAAGAGAGCA 2858  
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
DB 2859 GCTTATGAATCTTCAAGATAATTGATAAAGCAAGTATTGACAGCTATTTCGAAGAGT 2918  
QY 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
DB 2919 GGGCAAAACCAAGATAATTAAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2978  
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
DB 2979 CCATCTCGAAAGAGATTGAAGATCTTGAAGGCTGAACTGAAGTGCAGGTGGGCGAG 3038  
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400  
DB 3039 ACGGTGGCTGTTGGAAACAGTGGCTGTGGAGAGACACACAGTCCAGCTGATGCGAG 3098  
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
DB 3099 AGGCTCTATGACCCACAGAGGGGATGGTCACTGTTGATGGACAGCATATTAGGACATA 3158  
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
DB 3159 AATGTAAGTCTTCTCGGGAATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3218  
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
DB 3219 ACCAGATAGCTGAAACATTTGCTATGCGCGTGAATGTCAACATGATGATGATGATG 3278  
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480  
DB 3279 AAGCTGTCAAG 3338  
QY 481 ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500  
DB 3339 ACCTGTTGGAAG 3398  
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
DB 3399 GCACGTGCTGCTGGTTCGCAACCCCAAGATCTCTCTGATGAGGCCAGCTCAGCTTG 3458  
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534  
DB 3459 GACACAGAAAGCAAGCAGTGGTTCAGGTGGCTCTGGATAAG 3500

## RESULT 8

US-08-784-649A-1  
; Sequence 1, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,649A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: Reg.No. 5830697 36,677  
REFERENCE/DOCKET NUMBER: 06037/007001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4264 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-784-649A-1

Alignment Scores:  
Pred. No.: 1,31e-171 Length: 4264  
Score: 1545.00 Matches: 306  
Percent Similarity: 73.29% Conservative: 100  
Best Local Similarity: 55.23% Mismatches: 98  
Query Match: 56.41% Indels: 50  
DB: 2 Gaps: 7

US-09-873-409-7 (1-541) x US-08-784-649A-1 (1-4264)

```
QY 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 292 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATCGT 351
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 352 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAGAAATTTAGAGATCTGATGCA 411
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 412 AACATCTAATAGAGTATGATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 471
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 472 ATGACAGGATGCTATTATTACAGTGAATGGTCTGGGGTGGTGGTCTGCTTAC 531
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 532 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 592 TTTTTCATGCTATATACGACAGAGATAGCTGGTTGATGTCACGATGTTGGGGAG 651
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 652 CTTAACACCCGACCTTACAGATGATGCTCCAGAAATTAATGAGAAATTTGGACAAATTT 711
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 712 GGA---ATGTTTCATGCTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 768
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 769 GGTGGGAAGCTTAACCTTGTGATTTGGCCATCATGCTGTTCTTGGACGTGTCAGCTGCT 828
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 829 GTCTGGGCAAGATACCTATCTCTTACTGATTAAGAACTCTTAGCGGTATGCAAAAGCT 888
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
```

```
DB 889 GGAGCAGTAGCTGACAGAGCTTTGGCAGGCAATTAGAACCTGTGATTCATTTGGAGGACAA 948
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 949 AAGAAAGACTTTGAA----- 963
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 964 -----AGGTACACACAAAATTTAGAGAA 987
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 988 GCTAAAAGAAATTTGGGATAAGAAAGCTATTACAGCAATATTTCTATAGTGTCTGCTTTC 1047
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
DB 1048 CTGCTGATCATGCTATCTGCTGGCTTCTGGTATGGGACCACTTGGTCTCTCTCA 1107
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 1108 GGGGA-----TATCTATTGGACAAGTACTCACTGTATTCTTTCTGTATTAAATGGG 1161
QY 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1162 GCTTTTAGTGTGGACAGGCATCTCCAGCATTTGAAGCAATTTGCAAAATGCAAGAGGCA 1221
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1222 GCTTATGAATCTTCAAGATAATTGATAATAGCCAAATTTGACAGCTATTGCAAGAGT 1281
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1282 GGGCACAAACACGATAATATTAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1341
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1342 CCATCTCGAAAAGAAATTAAGATCTTTGAAGGGCTCAACCTGAAGGTGCAGAGTGGCAG 1401
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGln 400
DB 1402 ACCTGGCTGCTGTTGGAAACAGTGGCTGGGAAAGACACACAGCTCCAGCTGATGCAG 1461
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1462 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACCAT 1521
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1522 AATGTAAGTGTCTACGGGAAATCATTTGGTGTGGTGCAGTCAGGAACCTGTATTGTTGCC 1581
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluMetGlu 460
DB 1582 ACCAGATAGCTGAAACCATTCGCTATGGCCGGAATAATGTCACCATGGATGAGATGAG 1641
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1642 AAAGCTGTCAGAGCAACCATGCTATGATCTTATCATGAACCTGCCTCATAAATTTGAC 1701
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
DB 1702 ACCCTGTTGGAGAGAGGGGCCCAATGAGTGGTGGGAGAGAGAGATTCGCCATT 1761
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
DB 1762 GCACGTGCCCTGTTTGCRAACCCCAAGATCTCTGCTGGATGAGGCCAGCTCAGCCTTG 1821
QY 521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
DB 1822 GACACAGAAAGCAAGCAGTGGTTTCAAGTGGCTCTCGATAAG 1863
```

RESULT 9

US-08-784-649A-5

; Sequence 5, Application US/08784649A

Patent No. 5830697  
 GENERAL INFORMATION:  
 APPLICANT: Sikic, Branimir I  
 APPLICANT: Chen, Gang  
 TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
 TITLE OF INVENTION: CYCLOSPORIN MODULATION  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson  
 STREET: 2200 Sand Hill Road  
 CITY: Menlo Park  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94025  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/784,649A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sherwood, Pamela J  
 REGISTRATION NUMBER: Reg. No. 5830697 36,677  
 REFERENCE/DOCKET NUMBER: 06037/007001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-322-5070  
 TELEFAX: 415-854-0875  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4264 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-784-649A-5

Alignment Scores:  
 Pred. No.: 1,31e-171 Length: 4264  
 Score: 1545.00 Matches: 306  
 Percent Similarity: 73.29% Conservative: 100  
 Best Local Similarity: 55.23% Mismatches: 98  
 Query Match: 56.41% Indels: 50  
 DB: 2 Gaps: 7

US-09-873-409-7 (1-541) x US-08-784-649A-5 (1-4264)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAlaCysLeuProLeuMetProLeu 20  
 Db 292 ATGGTGGTGGAACTTTGGCTGCCATCCATCGGGCTGACTTCTCTCATGATGCTG 351  
 Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 352 GTGTTGGGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGTC 411  
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 412 AACATCACTAATAGAGTGTATCAATGATACAGGGTTCTTCATGATCTGGAGGAGAC 471  
 Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaLeuIlePheGlyTyr 61  
 Db 472 ATGACCAGGTATGCCATTATTACAGTGAATTTGGTGGGTGCTGGTGGTCTGCTTAC 531  
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 532 ATTCAGGTTTCATTTGGTGGCTGGCAGTGAAGACAAATACACAAATTTAGAAACAG 591  
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
 Db 592 TTTTTCATGCTAATAGCAGAGGAGATAGGCTGTTGATGTCACGATGTTGGGAG 651

Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 652 CTTAACACCCGAGTTACAGATGATGTCTCCAAAGATTAAAGGAATGGTGCAAAAT 711  
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 712 GGA--ATGTTCCAGTCATGCGACATTTTTCACCTGGGTTTATAGTAGGATTACACGT 768  
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 769 GGTGGAGCTAAACCTTGATTTTGGCCATCAGTCTGTTCTTGGACTGTGCTGCTGCT 828  
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 829 GTCTGGGCAAGATACATCTTCTTACTGATAAAGAACTCTTACGATGTCANAAAGCT 888  
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 889 GGAGCAGTAGCTGAGAGGTCTTGGCAGCAATTAGACCTGTGATTCATTTGGAGGACAA 948  
 Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
 Db 949 AAGAAAGAACTTGAA----- 963  
 Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
 Db 964 -----AGGTACACACAAATAATTTAGAGAA 987  
 Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
 Db 988 GCTAAAGAAATTTGGGATAAAGAGCTATTACAGCCAAATTTCTATAGTGTGCTTTC 1047  
 Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
 Db 1048 CTGCTGATCTATGATCTTATGCTCTGGCTTCTGCTATGGGACCACCTTGTGCTCTCA 1107  
 Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
 Db 1108 GGGGA-----TATTCTATTGGACAAGTACTCCTGTTATTTCTGTTAATTTGGG 1161  
 Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyVala 320  
 Db 1162 GCTTTTAGTGTGGACAGCATCTCCAGCATTGAAGCAATTTGCAATGCAAGAGGAGCA 1221  
 Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
 Db 1222 GCTTATGAATCTTCAAGATAATTGATATAAGCCAAGTATTGACAGCTATTTCGAAGAGT 1281  
 Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
 Db 1282 GGGCACAAACAGATAATATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTTCAGTTAC 1341  
 Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
 Db 1342 CCATCTCGAAAAGAGTTAAGATCTTGAAGGGCTTGAACCTGAAGGTGAGAGTGGGCAG 1401  
 Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGln 400  
 Db 1402 ACGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAGAGACACACAGCTCCAGCTGATGACG 1461  
 Qy 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420  
 Db 1462 AGGCTCTATGACCCACAGAGGGGATGGTCAGTGTGTGATGGACAGGATATTAGGACATA 1521  
 Qy 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440  
 Db 1522 AATGTAAGGTTTCTACGGGAAATCATTTGTTGGTGGTGAAGTCAAGAACCTGTATTGTTGCC 1581  
 Qy 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGlu 460  
 Db 1582 ACCAGATAGCTGAAACACATTGCTATGGCCGTGAAATGTCAACCATGATGATGATGAG 1641  
 Qy 461 ArgAlaAlaArgGluAlaAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480

```

1642 AAAGCTGTCAAGGAAGCAATGCTATGACTTTATCATGAAACTGGCTCATAAATTGAC 1701
481 ThrLeuValGlyGluLysGlyAlaGlnMetSerClyGlyGlnLysGlnArgIleAlaIle 500
1702 ACCCTGGTGGAGAGAGGGGCCAGTTGAGTGGTGGGAGAGCAGAGATCGCCATT 1761
501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuLeuAspGluAlaThrSerAlaLeu 520
1762 GCACGTGCCCTGGTTGCAACCCCAAGATCTCTGCTGGATGAGGCCAGCTCAGCCTTG 1821
521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
1822 GACACAGAAAGCAAGCAGTGGTTTCAAGTGGCTCTGGATAAG 1863

RESULT 10
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION NUMBER: 892,575
; APPLICATION DATA:
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3
```

```

Alignment Scores:
Pred. No.: 1,75e-171 Length: 4669
Score: 1544.50 Matches: 305
Percent Similarity: 72.92% Conservative: 99
Best Local Similarity: 55.05% Mismatches: 101
Query Match: 56.39% Indels: 49
DB: 6 Gaps: 6

US-09-873-409-7 (1-541) x 5206352-3 (1-4669)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGGTGGGAACCTTTGGCTGCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerApp-----AsnLeuIleSer 31
DB 635 GTGTTGGAGAAATGACAGATACTTTGGCAATGCAGGAATTTAGAAATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTy-SerPhePhe----- 43
DB 695 AACATCACTAATAGAGTGATATCAATGATACAGGGTTCCTTCATGATCTGGAGGAGAC 754
QY 44 -----ArgLeuThrLeuTyTyTyValGlyIleGlyValAlaLeuIlePheGlyTy 61
DB 755 ATGACCAAGGTATGCCTATTATTACAGTGGAAATTTGGTCTGGGGTGGTGTGCTGTAC 814
QY 62 IleGlnIleSerLeuTyrIleIleAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTTTGGTGGCTGGAGCTGGAGACAAATACACAAATTTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGTATGTCACGATGTTGGGAG 934
QY 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAAACACCGGACTTACAGATGATGTCTCTAAGATTAATCAAGATTATTTGGTGACAAAT 994
```

```

QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCAGTGGGTTTATAGTAGGATTTACACGT 1054
QY 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGTGAAGCTAACCCCTTGTGATTTTGGCCATCAGTCCCTGTTCTTGGACTGTCAGCTGT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTySerLysAla 180
DB 1115 GTCGGGCAAGATACATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGAGCAGTAGCTGAAGAGCTCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
DB 1235 AAGAAAGAACTTGAA----- 1249
QY 221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
DB 1250 -----AGGTACACAAATAATTTAGAGAA 1273
QY 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
DB 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTGCTGTTT 1333
QY 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn 280
DB 1334 CTGCTGATCTATGTCATCTTATGCTCTGGCTCTTGTGTATGGGACCACTTGGTCTCTCA 1393
QY 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
DB 1394 GGGGAA-----TATCTATTTGGCAAGTACTCCTGTTATTTCTGTATTAATTTGGG 1447
QY 301 SerTy-CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
DB 1448 GCTTTTATGTTGGACAGGCACTCTCCAAAGCATTTGCAAAATGCAAGAGGAGCA 1507
QY 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
DB 1508 GCTTATGAAATCTTCAAGATAATTTGATAAAGCCAAAGTATTGACAGCTATTTCGAAGAGT 1567
QY 341 GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr 360
DB 1568 GGGCAAAACCAAGATATATTAAGGAAATTTGGAATTTGAGAAATGTTCACTTCAGTTAC 1627
QY 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
DB 1628 CCATCTCGAAAGAAAGTTAAGATCTTGAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAG 1687
QY 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
DB 1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACACACAGTCCAGCTGATGTCAG 1747
QY 401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
DB 1748 AGGCTCTAGACCCCAAGAGGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1807
QY 421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
DB 1808 AATGTAAGGTTTCTACGGGAATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1867
QY 441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluMetGlu 460
DB 1868 ACCACGATAGCTGAAACATTCGCTATGCGCGTGAATAATGTCCACCATGATGATGATGATG 1927
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
DB 1928 AAGCTGTCAAGGAAGCAATGCTATGATTTTATCATGAACTGCTCATAAATTGAC 1987
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
```

Db 1988 ACCCTGGTGGAGAGAGGGCCCGAGTTGAGTGGGCGACAGAGGATCGCCATT 2047  
Qy 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520  
Db 2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCTCTGCTGGATGAGGCCACGTCAGCCTTG 2107  
Qy 521 AspSerGluSerLysSerAlaValGlnAlaLeuLeuGluLys 534  
Db 2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149

## RESULT 11

US-08-583-276-18  
; Sequence 18, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:  
; APPLICANT: McDonagh, Kevin T.  
; APPLICANT: Nienhuis, Arthur  
; APPLICANT: Tolstoshev, Paul  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carelia, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi & Stewart  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: DW4.V2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,276  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/332,444  
; FILING DATE: 31-OCT-1994  
; APPLICATION NUMBER: 07/887,712  
; FILING DATE: 22-MAY-1992  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4669 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: singular  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: Genomic DNA

US-08-583-276-18  
Alignment Scores:  
Pred. No.: 3,066-169 Length: 4669  
Score: 1525.50 Matches: 301  
Percent Similarity: 72.56% Conservative: 101  
Best Local Similarity: 54.33% Mismatches: 103  
Query Match: 55.70% Indels: 49  
DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-583-276-18 (1-4669)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 575 ATGGTGGTGGAACTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 635 GTCTTGGAGAAATGACAGATATCTTTGCAATGCAGGAAATTTAGAGATCTCATGATCA 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 695 AACATCACTAATAGAAAGTGATATCAATGATACAGAGGTTCTTCAATGAATCGAGGAGAC 754  
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 755 ATGACACAGGTATGCTATATTACAGTGGAAATGGTGGGGTGGTGGTGGTGGTGGTGGTGG 814  
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 815 ATTCAAGTTTCAATTTGGTGGCTGGCAGCTGGAACAATAACACAAATATTAGAAACAG 874  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 875 TTTTTCATGCTAATACGACAGAGATAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 934  
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 935 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAAAGAACTTATTGGTGCACAAAT 994  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTCTACCTGGGTTTATAGTAGGATTACACGT 1054  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 1055 GGTGGAGCTAACCCCTTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGTGTCT 1114  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 1115 GTCTGGCAAGATATCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCCAAGCT 1174  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 1175 GGAGCAGTAGCTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTGGAGACAA 1234  
Qy 201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220  
Db 1235 AAGAAAGAACTTGAA----- 1249  
Qy 221 ProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyrThrGlnAsnLeuLysAsp 240  
Db 1250 -----AGGTACACAAATAATTTAGAGAA 1273  
Qy 241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260  
Db 1274 GCTAAAGAAATTTGGGATAAAGAAAGCTATTACAGCAATATTTCTATAGTGTGCTTTC 1333  
Qy 261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280  
Db 1334 CTGCTGATCTATGATCTTATGCTTGGCTTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCA 1393  
Qy 281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300  
Db 1394 GGGGAA-----TATTCATTGGACAAGTACTCCTGTTTCTTTCTGTTAATTTGGG 1447  
Qy 301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320  
Db 1448 GCTTTTAGTGGTGGACAGCACTCCAGCATTTGAAGCATTTGCAATGCAAGAGGCA 1507  
Qy 321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340  
Db 1508 GCTTATGAAATCTTCAAGATAATTGATAATAGCCCAAGTATTGACAGCTATTGCAAGAGT 1567  
Qy 341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360  
Db 1568 GGGCACAAACCAAGATAATTAAGGGAAATTTGGAATTCAGAAATTTGTCATTTCAGTTAC 1627  
Qy 361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380  
Db 1628 CCATCTCGAAAGAAAGTGAATCTTGAAGGGCTGAACTGAGGTGAGAGTGGGAG 1687  
Qy 381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400

Db	1688	ACGGTC	CCCTCG	TGGTGG	AACAGT	GCGCTGTG	GGGAAG	ACCAAC	AGTCC	ACAGTCC	ACGTGAT	GCAG	1747
Qy	401	ArgLeu	TyrAsp	ProAsp	AspGly	PhelIle	MetVal	AspGlu	AsnAsp	IleArg	AlaLeu	420	
Db	1748	AGGCTC	TATGAC	CCCCAC	AGAGGG	GATGGT	CAGTGT	TCATG	GCACAG	GAATATT	TAGAG	ACCATA	1807
Qy	421	AsnVal	ArgHis	TyrArg	AspHis	IleGly	ValVal	SerGln	Glupro	ValLeu	PheGly	440	
Db	1808	AATGTA	AGGTTT	CTACGG	GAATCAT	TGTGTG	TGTGAGT	CAGAA	CCCTGAT	TGTTG	TGTC	1867	
Qy	441	ThrThr	IleSer	IleAsn	IleLys	TyrGly	ArgAsp	AspVal	ThrAsp	GluGlu	MetGlu	460	
Db	1868	ACCACG	TAGTCT	GAACAT	CTCGTAT	GCCGCTG	GAATAAT	GTCCAC	CATGGAT	GAGATT	GCAG	1927	
Qy	461	ArgAla	IleArg	GluAla	IleAla	TyrAsp	PhelIle	MetGlu	PhePro	AsnLys	PheAsn	480	
Db	1928	AAAGCT	GTCGAA	GGAAGC	AATGCT	TATGACTT	TATCAT	GAAACT	CGCTCAT	TAAATTT	GCAC	1987	
Qy	481	ThrLeu	ValGly	GluLys	GlyAla	GlnMet	SerGly	GlyGln	LysGln	ArgIle	AlaIle	500	
Db	1988	ACCTCG	TGTTGG	AGAGAG	GAGGCC	CCAGTTG	AGTGTG	TGGCAG	ACAGAC	GAGATGC	GCATT	2047	
Qy	501	AlaArg	AlaLeu	ValArg	AsnPro	LysIle	IleLeu	IleLeu	AspGlu	AlaThr	SerAla	Leu	520
Db	2048	GCACGT	CCCTCG	TGTTGC	ACACCA	CCCAAGAT	CTCCTCG	TGATG	AGGCC	ACCGTC	AGCCTTG	2107	
Qy	521	AspSer	GluSer	IleLys	SerAla	ValGln	AlaLeu	GluLys	534				
Db	2108	GACACG	AAAGC	GAAGCAG	TGGTGT	TCAGGT	CGGCTCG	GATTAAG	2149				

RESULT 12

```

US-08-612-521-1
; Sequence 1, Application US/08612521
; Patent No. 5786463
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B
; APPLICANT: Skatrud, Paul L
; APPLICANT: Thornewell, Susan J
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
; TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Division/AEH
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,521
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamilton, Amy E
; REGISTRATION NUMBER: 33,894
; REFERENCE/DOCKET NUMBER: X-9693
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3169
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

```

```

; LOCATION: 1..4224
US-08-612-521-1

Alignment Scores:
Pred. No.: 1,47e-102 Length: 4224
Score: 960.00 Matches: 210
Percent Similarity: 56.89% Conservative: 112
Best Local Similarity: 37.10% Mismatches: 182
Query Match: 35.05% Indels: 62
DB: 1 Gaps: 8

US-09-873-409-7 (1-541) x US-08-612-521-1 (1-4224)
Qy 1 MetlleuGlylleuAlaSerleuValAenGlyAlaCysLeuProLeuMetProleu 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 466 ATGTCCTTGGCGCTTGTGCTTCGCTGCAGCCGAGTCTGCCAACCTTTAATGACTTTA 525
Qy 21 ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeuValGlnThrAsnThr-Tyr 40
::: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 526 ATATTCGTCGTATACCAAGCTCTTTCAGCAATATGCTGTCAT-T---GCCAACCAATA 582
Qy 41 SerPhePheArgLeuThr----- 46
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 583 TCCCAAGGGGACTTACTCCGAGAGCGTCTGGCGCTTTACAAGCAGCCAAAGATGATCTC 642
Qy 47 -----leuTyrTyrValGlylleGlyValAlaAleuIle 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 643 AAGACTCAATCTGCCACACAATGCCCTTTATCTCGATGGCCATTGGCAATTTGGAATTTCTA 702
Qy 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 703 CGGACTTGGCTTTACAGTTCATCTGGAATGTACTCGGAGCTCACTCTAAAGGATT 762
Qy 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 763 AGAGAGCGTTACTTGGCTGCAGTGCTTAGACAGGAGATTGCTCTTTGATGATCTGGGC 822
Qy 99 IleGlyGluLeuAsnThrArgMet---ThrAspIleAspLysIleSerAspGlylleGly 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 823 GCCGGAGAACTCGCCACTCGCATTCAGACCGGATTCGTCACCTGTGTCCAAGAGGGAACATCT 882
Qy 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 883 GAAAGTTCGCTCTGATATCCAAATATGCTGTACTTTGTGCGGTTTGTCTAGCT 942
Qy 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 943 TTTGTCGGTCACTCGTCTTTGGCGCGCTCTTGTTTCCATCTGCCAGTCATCATGCTC 1002
Qy 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1003 TCGCGCGGTATTATGATGACGGCCATATGGCCAAATATGGAGACTCGGCTCTTGATCACATC 1062
Qy 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
::: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 1063 CCGAAACGGGCGAGCTTGGCAGAGAGGTCATAGGAAGTATCAGGACTGTTTCAGGCTTTT 1122
Qy 198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr-AlaTrp 217
::: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 1123 --GCCAAGGAAAAGATCTTTGGTGACAAATTT----- 1152
Qy 218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
::: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 1153 -----GCCGATCAC 1161
Qy 238 LeuLysAspAlaLysAspPheGlylleLysArgThrIleAlaSerLysValSerLeuGly 257
::: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 1162 ATTGAGCAGCAGAGATTGTCGTAGAAAGGCTCCATCTTTGAAGGCTTTGGTTTGAGC 1221
Qy 258 AlaValTyrPhePheMetAenGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
::: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 1222 ATCATGTTCTTGTCATCTACGCGGCTTATGCCCTCGCCTCTTCTTACGGTGCAATCTC 1281

```



```
QY 278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db 1282 GTTCAGTAAAGCGCCAGCC-----GACTCCGGTATCGTCAACAGCTTTTCATGTCCATC 1335
QY 298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db 1336 CTCATTGGTTCCTTCCATGGCGATGCTGTGCCGAGTGGCTGCTGTCCACCAAGCG 1395
QY 318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspIleAspIleAspIleAspPhe 337
Db 1396 CGAGGCGCTGCTCCCAAGCTTTTCGCAACTATCGACCGTGTACCCGCCATTGATCCGCC 1455
QY 338 SerThrAlaGlyTyrIleProGluSerIleGluGlyThrValGluPheIleAsnValSer 357
Db 1456 AGCGAGGAAGGTTTCAAGCCCGATGCGCTTTCGCAAGTATGAGTTCGAAATGTCAG 1515
QY 358 PheAsnTyrProSerArgProSerIleIleLeuIleGlyGlyLeuAsnLeuArgIleLys 377
Db 1516 TTCATTATCCGTCGCCAGCCAGCATCTCTATCTTGAAGGCTTCACTACTAGTTTGA 1575
QY 378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyIleSerThrValValGln 397
Db 1576 GCGGTAAGACTTTCGCTCTCGTCGAGCTAGTGAAGCGGAAGAGTACTGCTGCTTCT 1635
QY 398 LeuLeuGluArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db 1636 CTTATTGAGAGGTTTACGACCCCGTTAGTGGTGTGTGAAGCTTGACCGTAGAGATATC 1695
QY 418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db 1696 AGTCCGCTTAACCTCAATGGCTTGTCAACAGATTGGTCTTGTATCTCAGGAACCACT 1755
QY 438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly-----ArgAsp 452
Db 1756 CTTTTCGGTACCACCGTCCGCGGTACGTCGAACACCGTCTCTCGCTCCACATATGAG 1815
QY 453 AspValThrAspGluGlu-----MetGluArgAlaAlaArgGluAlaAsnAla 468
Db 1816 AATGCCTCTCTCAGGAAATAATTGAGCTTGTGAAGAGGCTTGTGTCGATGCCAAGCC 1875
QY 469 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 488
Db 1876 CATAATTTTATCATGAACACTTCCGCAAGGTATGATACCATGGTGTGAGCGCGGTATG 1935
QY 489 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 508
Db 1936 CTGTTATCTGGTGTGAGAGCAGCGAGTTCGATTCGCTATCGTTTCGACCT 1995
QY 509 LysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
Db 1996 AGGATTTTGTGTGACGAAAGCTACCTCTGCTCTTGACACTCAGACGGAAGTATTGTG 2055
QY 529 GlnAlaAlaLeuGluLys 534
Db 2056 CAAGATGCTTTGGACAG 2073

RESULT 13
US-08-612-734B-1
; Sequence 1, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Eli Lilly and Company
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
```

```
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4047 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4047
US-08-612-734B-1

Alignment Scores:
Pred. No.: 1,07e-99 Length: 4047
Score: 935.50 Matches: 203
Percent Similarity: 57.01% Conservative: 110
Best Local Similarity: 36.98% Mismatches: 191
Query Match: 34.15% Indels: 45
DB: 2 Gaps: 7

US-09-873-409-7 (1-541) x US-08-612-734B-1 (1-4047)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 331 ATACTCGTGAGCGCAATCTGCTATGCTGTGGTGGCGCTCTTCCCTTGTTCACAAAT 390
QY 21 ValLeuGlyGluMetSerAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 391 CTCTTCGGTTCAGTACC---TCAGGGTTCAGGCGATATCTCTGGGTACTATGCGCTAT 447
QY 41 SerPhePhe-----ArgLeuThr-----LeuTyrTyrValGlyIleGlyValAla 55
Db 448 CATGAGTTCTATCACAACACTGACTAAGAAATGCTTTACTTTGTGTATCTCGGTATTGCC 507
QY 56 AlaLeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThr 75
Db 508 GAGTTTGCACAGTCTATGTCAGCCCGGGTTCATTTTACTCGCGACATCTCACA 567
QY 76 LysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAsp 95
Db 568 CAGAAGATCCGTGAAATATCTTGAGGCTATCTCAGGCGCAAGATATGCTTACTTCGAC 627
QY 96 SerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAsp 114
Db 628 AAGTTGGCGCGCGGTGAAGATTACACCGCTATCAGTGTGATACCAACCTGATCCAGGAC 687
QY 115 GlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeu 134
Db 688 GCATCTCTGAGAAAGTTGGTCTCACTTTGACCGGATTCGCCACATTTGTAACCGCATTT 747
QY 135 AlaValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeu 154
Db 748 ATGTGCGCTACGTCAAGTATGGAAGTTGGCTCTGATCTGTACCTCACTATCGTCGCG 807
QY 155 IleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeu 174
DB: 174
```

Db 808 CTGGTCATGTTATGGAGGTGGGTGAGGTTATTTGTGAAGTACGACCAAGAAATCTATT 867  
Qy 175 SerAlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrVal 194  
Db 868 GAAAGTTATGGTCTGCTGAGTCTGCGGAGAGTCTATCAGTCCATTCGGAATGCT 927  
Qy 195 IleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAenIleThrArg 214  
Db 928 ACCGTTTCGGCACTCAGGATAAGTCGCCAAG----- 960  
Qy 215 TyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu\*\*\*PheValArgTyr 234  
Db 961 -----CAATAC 966  
Qy 235 ThrGlnAenLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysVal 254  
Db 967 GAAACCCATCTGGCTGAGGTGAAATAATGGGCGTCAACACAGGTCACTCTGGTGTATG 1026  
Qy 255 SerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGly 274  
Db 1027 ATGATTGGTGTATGTTGCGTATCATGTTCTCGAAGTATGTTCTCGGTTTCTGGATGGGA 1086  
Qy 275 ThrSerLeuIleLeuAenGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhe 294  
Db 1087 TCTCGCTTCGTTGCGTAAGAA-----GTCAACGTGGGCCCAAGTCTGACAGTTTTC 1140  
Qy 295 PheSerValIleHisSerTyrCysIleGlyAlaValProHisPheGluThrPhe 314  
Db 1141 ATGCTATCTGATCGTTGTTTCAGTTTGGGCACGTCCGCCCAATGTCAGGCCCTTT 1200  
Qy 315 AlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIle 334  
Db 1201 ACGAATGGTGTGTCGCGCGCGGAAGATTACAGCAGGATTGACGCAGATCGCCACTG 1260  
Qy 335 AspAenPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLys 354  
Db 1261 GACCCCTATTCTGACGAAGGAAGGTACTACACATTTTGAAGGAATATCGAATTCGC 1320  
Qy 355 AenValSerPheAenTyrProSerArgProSerIleLysIleLysGlyLeuAenLeu 374  
Db 1321 AATGTCAAACACATCTACCTTCAAGACCCGAGTTACAGTACGGAAGATGTCCTTTA 1380  
Qy 375 ArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAenGlySerGlyLysSerThr 394  
Db 1381 TCGATGCCCGCGGAAGACTACCGCATTTGTTGGCCCATCTGGCTCTGGAAGAGTACT 1440  
Qy 395 ValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGlu 414  
Db 1441 GTTCTCGGCTTGGTGGAGCGCTTTTACCTTCCAGTAGGAGGCCAGGTATTGCTGGACGGC 1500  
Qy 415 AsnAspIleArgAlaLeuAenValArgHisTyrArgAspHisIleGlyValValSerGln 434  
Db 1501 CATGATATCAAAACCCCTCAACCTCGTGGTGGTGGCGACAGCATCTCTCTTGTGAGCCAG 1560  
Qy 435 GluProValLeuPheGlyThrThrIleSerAsnAenIleLysTyrGly----- 450  
Db 1561 GAACCTGTTCTTTTCAGCACCATCTTTAGAAACATCGACATGCTGATTGGCCACC 1620  
Qy 451 -----ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465  
Db 1621 AAATTCGACGATGATGTCGAGGACGAAGATTAGAGAGTCTGTTGAGAATGCGGCCAGAA 1680  
Qy 466 AlaAenAlaTyrAspPheIleMetGluPheProAenLysPheAenThrLeuValGlyGlu 485  
Db 1681 GCCAATGCTCATGATTTTATTTATGGCTCTGCTGAAGTTACGATACGAATGTTGGGTCAG 1740  
Qy 486 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505  
Db 1741 CGTGGTTCTTACTTTTCAGGAGGTGAGAAGCAAGTATGCTGCTGCGCCATTGTC 1800  
Qy 506 ArgAenProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525  
Db 1801 AGTGACCCCAAGATTCTGTTGATGAAGCTACATCAGCTTTGGGATACCAAGTCCGAG 1860

Qy 526 SerAlaValGlnAlaAlaLeuGluLys 534  
Db 1861 GCGCTGCTCAAGCGCTCTTGATAAA 1887  
RESULT 14  
US-08-612-734B-3  
; Sequence 3, Application US/08612734B  
; Patent No. 5914246  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: Tobin, Matthew B.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene of  
; TITLE OF INVENTION: Aspergillus fumigatus  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center, DC1501  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.A.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,734B  
; FILING DATE: 08-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Craig, Anne I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: X-9681  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-354-9570  
; TELEFAX: 617-354-4043  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4800 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
US-08-612-734B-3  
Alignment Scores:  
Pred. No.: 2,86e-97 Length: 4800  
Score: 916.00 Matches: 202  
Percent Similarity: 55.47% Conservative: 102  
Best Local Similarity: 36.86% Mismatches: 182  
Query Match: 33.44% Indels: 62  
DB: 2 Gaps: 7  
US-09-873-409-7 (1-541) x US-08-612-734B-3 (1-4800)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 693 ATTCTCTCGGTTTCAGTACGCTTCCAGGCGATATCTCTGGGTACTATGCCCTAT 752  
Qy 21 ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 753 -----CAT 755  
Qy 41 SerPhePhe---ArgLeuThr-----LeuTyrTyrValGlyIleGlyValAlaAla 56  
Db 756 GAGTTCATCAACAACCTGACTAAGAATGTGCTTTACTTTGTGTATCTCGGTATTGCCGAG 815  
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76  
Db 816 TTTGTACAGTCTATGTCACGACCGTGGGTTTCATTTATATTACTGGCGAATCTCACA 875

Qy	436	ProValLeuPheGlyThrThrIleSerAsnAniLeLysTyrGly	-----	450
Db	1869	CCTGTTCTTTTCAGCACACGATCTTTAGAAACATCGAACATGGCTTGATGGCACCAAA	1928	
Qy	451	-----ArgAspAspValThrAspGluGluMetGluAtrGAlaAaArgGluAla	466	
Db	1929	TTGAGCATGATCGAAGGACAGATTACAGAGCTCGTTGAGAAATCGCGCCAGATGCGC	1988	
Qy	467	AsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys	486	
Db	1989	AATGCTCATGATTTATTATGCTCTGCTCCCTGAAGGTTACGATACCGAATGTGGTCCAGCGT	2048	
Qy	487	GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAaArgAlaLeuValArg	506	
Db	2049	GGTTTCTTACTTTCAGAGGCTCAGAACACGTATTGCCATTGCTCGTGCCATTGTCAGT	2108	
Qy	507	AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer	526	
Db	2109	GACCCCAAGATTCGTGCTTGATGAAGCTACATCAGCTTTGGATACCAAGTCCGAGGCG	2168	
Qy	527	AlaValGlnAlaAlaLeuGluLys	534	
Db	2169	GTGCTCCAAGCGCTCTTGATAAA	2192	
RESULT 15				
US-08-996-545-1				
; Sequence 1, Application US/08996545				
; Patent No. 592898				
; GENERAL INFORMATION:				
; APPLICANT: Skatrud, Paul L.				
; APPLICANT: de Waard, Maarten A.				
; APPLICANT: Peery, Robert B.				
; APPLICANT: Andrade, Alan C.				
; TITLE OF INVENTION: Multiple Drug Resistance Gene at rD of				
; TITLE OF INVENTION: Aspergillus nidulans				
; NUMBER OF SEQUENCES: 3				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Eli Lilly and Company				
; STREET: Lilly Corporate Center				
; CITY: Indianapolis				
; STATE: Indiana				
; COUNTRY: U.S.				
; ZIP: 46285				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: Patent In Release #1.0, Version #1.30				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/996,545				
; FILING DATE:				
; CLASSIFICATION: 435				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Webster, Thomas D.				
; REGISTRATION NUMBER: 39,872				
; REFERENCE/DOCKET NUMBER: X-11766				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 317-276-3334				
; TELEFAX: 317-276-2763				
; INFORMATION FOR SEQ ID NO: 1:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 4002 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: cDNA				
; HYPOTHETICAL: NO				
; ANTI-SENSE: NO				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: 1..4002				
; US-08-996-545-1				

## Alignment Scores:

Pred. No.: 3,62e-97 Length: 4002  
 Score: 914.00 Matches: 201  
 Percent Similarity: 56.93% Conservative: 103  
 Best Local Similarity: 37.64% Mismatches: 181  
 Query Match: 33.37% Indels: 50  
 DB: 2 Gaps: 6

US-09-873-409-7 (1-541) x US-08-996-545-1 (1-4002)

```

QY 15 LeuProLeuMetProLeuValLeuGluMetSerAspAsnLeuLeuSerGlyCysLeu 34
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 368 GTGCCATTGTCGCGCTCGACCTTCAGAA- - - - -GCATAATGTTAT 409
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 35 ValGlnThrAsnThrTyrSerPhePhe- - - - -ArgLeuThrLeuTyrVal 50
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 410 ATCAAAATCTCGTACGAC -GAGTTCTATGATGAATGTGACCAAGACGTACTGTTCTGTA 468
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleThr 70
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 469 TACCTCGTATCGCGAGTTGTGTCATGCTCTATGTTAGTACTGTTGGCTTCATCTATACC 528
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 71 AlaAlaArgGlnThrIleArgIleArgGlnPhePheHisSerValLeuAlaGlnAsp 90
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 529 GGAGAACACGCCACGACAGAGATCCGCGAGTATTACCTTGAGTCTATCTCTGCGCCAGAAC 588
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 589 ATTGGCTATTGTTGATAAATCTCGGTGCGGGAGAGTGACCCCGTATTAACGCGGATACA 648
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 649 AACCTTATCCAGGATGCAATTCGGAGAGGTGGTCTCTACTTTGACTGCGCTGGCGACA 708
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 130 PheSerIleGlyLeuAlaValGlyLeuValIysGlyTrpLysLeuThrLeuValThrLeu 149
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 709 TTCGTGACAGCATTCATTACCTGCTCAATACCTGGAAGTTGGCTCTAAATTTGCAGC 768
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 769 TCACAAATTTGGCCCTCGTTCTCACCATGGCGGTGGTCTCAGTTTATCATCAAGTAC 828
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 829 AGCAAAAGTCGTGTGACAGCTAGTGGCGGCACTGTTGGGAAGAGTCAATCAGC 888
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 889 TCCATCAGAAATGCCACAGCGTTTGGCACCCACAGCAAGCTTGGCAAG- - - - - 936
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 210 LeuAsnIleThrArgTyrAlaTyrPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 936 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 936
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 937 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 937
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 988 ATTGTCATGGGTTTCATGATGGCGCCATGTTGGCCTTATGTAATCGAACTACGGTCTT 1047
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1048 GGCTTCTCGATGGGTTCTCGTTTCTCGTATGATGGT- - - - -GCAGTGTGATGGGTGAT 1101
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1102 ATTCTCAGAGTTCTCATGGCCATCTTGATCGGATCGTTCTCTTGGGGAAGCTTAGTCCA 1161
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1162 AATGCTCAAGCATTTACAAACGCTTGGCGCGCGCGCAAGATATTTGGAAACGATCGAT 1221
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Search completed: March 31, 2003, 15:37:48  
 Job time : 148.227 secs

```

QY 330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1222 CGCCAGTCCCATTTAGATCCATATTCGAAGGAGGAGCGCTCGACCATTTTGAAGGC 1281
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1282 CACATTGAGTTACGCAATGTCAGCATATTTACCCATCTAGACCCCGAGGTCAACGTCATG 1341
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1342 GAGGATGTTTCTGTCAATGCCCGCTGGAACCAACCGCTTTAGTCGCCCTCTGGC 1401
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1402 TCTGGAAAAAGTACGCTGGCTGGTGTGAGCGATTCTACATGCTGTTCGCGTACG 1461
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1462 GTTTTGTGATGGCGCATGACATCAAGGACCTCAATCTCGCTGGCTTCGCCAACAGATC 1521
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1522 TCTTTGTTAGCCAGGAGCTGTTCTTTTGGCACCAGCATTTATAGAAATATTAGGCAC 1581
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 450 Gly- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 460
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1582 GGTCTCATCGGCACAAAGTACGAGTAATGTAATCCGAGTAAGTCCGGGAACATCATCGAG 1641
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1642 AACCGCGCAAAATGCGAATGCTCATGACTTTATTACTGCTTGCCTGAAGGTTATGAG 1701
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1702 ACCAATGTTGGCGAGCGTGGCTTCTCTTCAGTGGCGCAGAAACAGCGCATTTGCAATC 1761
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1762 GCCCGTGGCGTGTGTAGTACCCCAAAATCTCTGCTCTGTGATGAAGCTACTTCGGCCTTG 1821
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 521 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 534
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1822 GACACAAATCCGAGGCGTGGTTCAAGCAGCTTTTGGAGAGG 1863
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 13:52:37 ; Search time 187.092 Seconds  
(without alignments)  
2337.888 Million cell updates/sec

Title: US-09-873-409-8  
Perfect score: 2592  
Sequence: 1 MILGLASLVNAGCLPLMLP.....SEKSAVOALEKOTPRYSF 514

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame.p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115424\_19290/app\_query.fasta\_1.7544  
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09873409 @CGN 1 1 960 @runat\_27032003\_115424\_19290  
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2592	100.0	1940	US-09-873-409-16 Sequence 16, Appl
2	2568.5	99.1	2021	US-09-873-409-15 Sequence 15, Appl
3	2552	98.5	3621	US-09-873-409-14 Sequence 14, Appl
4	2528.5	97.6	3702	US-09-873-409-13 Sequence 13, Appl

5	1699.5	65.6	3177	10	US-09-873-409-12	Sequence 12, Appl
6	1590	61.3	4189	10	US-09-866-866A-5	Sequence 5, Appl
7	1579	60.9	3860	10	US-09-866-866A-1	Sequence 1, Appl
8	1576	60.8	4317	9	US-10-044-671-1	Sequence 1, Appl
9	1571.5	60.6	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
10	1570	60.6	3860	10	US-09-866-866A-3	Sequence 3, Appl
11	1570	60.6	8630	10	US-09-306-417-1	Sequence 1, Appl
12	1570	60.6	8630	10	US-09-306-417-2	Sequence 2, Appl
13	1566.5	60.4	4425	10	US-09-769-097-3	Sequence 3, Appl
14	1564.5	60.4	4369	10	US-09-769-097-1	Sequence 1, Appl
15	1562.5	60.3	4643	9	US-10-072-621-2	Sequence 2, Appl
16	1559.5	60.2	4788	10	US-09-866-866A-7	Sequence 7, Appl
17	1479	57.1	3924	10	US-09-880-107-2299	Sequence 2299, Ap
18	1445	55.7	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	1243.5	48.0	5010	10	US-09-917-800A-483	Sequence 483, App
20	1141.5	44.0	1175	10	US-09-873-409-11	Sequence 11, Appl
21	1101.5	42.5	2856	10	US-09-873-409-10	Sequence 10, Appl
22	1050.5	40.5	1810	10	US-09-749-340-4	Sequence 4, Appl
23	1050.5	40.5	4175	10	US-09-749-340-3	Sequence 3, Appl
24	1017	39.2	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
25	965.5	37.2	3861	9	US-09-938-842A-263	Sequence 263, App
26	937.5	36.2	4002	9	US-09-758-828-1	Sequence 1, Appl
27	937.5	36.2	4002	9	US-09-758-828-3	Sequence 3, Appl
28	915	35.3	4653	9	US-10-101-388-2	Sequence 2, Appl
29	828	31.9	2066	10	US-09-873-409-9	Sequence 9, Appl
30	809	31.2	3792	9	US-09-882-694-10	Sequence 10, Appl
31	766.5	29.6	8036	9	US-10-101-388-1	Sequence 1, Appl
32	755.5	29.1	2681	10	US-09-749-340-8	Sequence 8, Appl
33	744.5	28.7	3512	10	US-09-749-340-7	Sequence 7, Appl
34	733.5	28.5	2698	10	US-09-749-340-5	Sequence 5, Appl
35	712	27.5	3999	9	US-09-882-694-9	Sequence 9, Appl
36	668	25.8	2298	9	US-10-156-239-18	Sequence 18, Appl
37	668	25.8	3408	9	US-10-156-239-16	Sequence 16, Appl
38	668	25.8	3408	10	US-09-795-693-16	Sequence 16, Appl
39	668	25.8	3512	9	US-10-072-621-1	Sequence 1, Appl
40	619.5	23.9	2674	10	US-09-917-800A-1500	Sequence 1500, Ap
41	619.5	23.9	2674	10	US-09-815-242-9908	Sequence 9908, Ap
42	607	23.4	1749	10	US-09-815-242-9908	Sequence 40, Appl
43	603	23.3	2247	9	US-09-870-759-40	Sequence 6018, Ap
44	602	23.2	1749	10	US-09-815-242-6018	Sequence 170, App
45	598	23.1	2824	9	US-09-954-531-170	

ALIGNMENTS

RESULT 1  
US-09-873-409-16  
; Sequence 16, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 16  
; LENGTH: 1940  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-16

Alignment Scores:  
Pred. No.: 1.14e-267  
Score: 2592.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 10  
Matches: 1940  
Conservative: 514  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-873-409-8 (1-514) x US-09-873-409-16 (1-1940)

Qy	1	MetIleLeuGlyIleIleuAlaSerIleuValAanGlyAlaCysIleuProLeuMetProLeu	20
Db	34	ATGATCTCGGTATATCGGCATCACGTGTCAAAGAGAGCTGCCCTTCCTTTAATGCCACTG	93
Qy	21	ValLeuGlyGluMetSerAspAenLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	94	GTTTTAGGAGAAATGAGTGATACCTTATTAGTGGATGCTAGTCCAAACTAACACATAC	153
Qy	41	SerPhePheArgIleuThrIleuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly	60
Db	154	TCCTTCTTCAGGTGACCCCTGTATTATGTGTGAATAGGTGTGTGCTGCCCTTGATTTTGGT	213
Qy	61	TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysAaGileAaGlyLys	80
Db	214	TACATACAGATTTCCCTTGTGGATTAACTCGACGACGACAGACCAGAGAGATTTCGAAAA	273
Qy	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly	100
Db	274	CAGTTTTTTTCATTTCAGTTTTGGCACAGGACATCGCTGCTTTGATAGCTGTGACATCGGT	333
Qy	101	GluLeuAsnThrArgMetThrAspIleAspLysIleIleSerAspGlyIleGlyAspLysIle	120
Db	334	GAACCTTAACACTCGCATGCAGACATGTGACAAAATCAGTGAATGTTGGAGATAAGATT	393
Qy	121	AlaLeuLeuPheGlnAanMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	394	GCTCTGTGTGTTTCAAAACAATGCTACTTTTTTCGATTTGGCTGGCGATGGTTGGTGGAG	453
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	454	GGCTGGAAACTCACCTAGTGACTCTATCCACGCTCTCTCTTATAATGGCTTCAGCGGCA	513
Qy	161	AlaCysSerArgMetValIleSerIleuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	514	GCATGTCTTAGGATGGTCATCTCATTGACGATAGGAATTAAGTGGCCTATTTCCAAAGCT	573
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	574	GGGCGCTGTCGCAGAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTTAGGCCCCAG	633
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAanLeuLysAspAlaLysAspPheGlyLysLys	220
Db	634	GAGAAAGAATCTTCAAAAGGTATACACAGAATCTCAAAGATGCAAAAGGATTTTGGCATAAAA	693
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	694	AGGACTATAGCTTCAAAAGTGCTCTCTTGGTGTGTGTACTTCTTTATGAAATGGAACTAT	753
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAanGlyGluProGlyTyrThrIle	260
Db	754	GGACTCTCTTTTGGTATGGAACCTCTCTTGATTTTAATGGAGAACCTGGATATACCATC	813
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyValAla	280
Db	814	GGGACTGTCTCTGCTGTTTCTTTTAGTGTAAATCCATAGCAGTATTGTCAATTTGGAGACGA	873
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
Db	874	GTCCCTCACCTTTGNAACCTTCGCAATAGCCCGAGGAGCTGGCTTTCATATTTTCCAGTT	933
Qy	301	IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	934	ATTGATTAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCCTGGAATCCATA	993
Qy	321	GluGlyThrValGluPheLysAanValSerPheAsnTyrProSerAtqProSerIleLys	340
Db	994	GAAGAACTGTGGAAATTTAAAAATGTTTTCTTCAATATTCCAAATCAAGACCATCATCAAG	1053
Qy	341	IleLeuLysGlyLeuAanLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	1054	ATTCTGAAAGGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCCGCTCGCTCGGTCTC	1113

Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp	380
Db	1114	AATGGCAGTCGGGAAGTAGTACGGTAGTCCAGCTTCTCGACAGGTTATATGATCCGGATGAT	1175
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	1174	GGCTTTATCATGTGTGGATGGAATGACATCAGAGCTTTAAATGTGGCGCATTATCGAGAC	1233
Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle	420
Db	1234	CATATTCCGAGTGGTGTAGTCAAGAGCCGTGTTGTTCTGGGACCACCATCAGTAACCAATATC	1299
Qy	421	LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn	440
Db	1294	AAGTATCGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGAGAACCAAT	1355
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1354	CGGTATCATTTTATCATGGAGTTTCTTAATAAATTTAATACATTGGTAGGGGAAAAAGGA	1413
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1414	GCTCAAAATGAGTCGGGGCGAGAACAGAGGATCGCAATTTGCTCGTGCTTACTTTCGAAC	1473
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1474	CCCAGAATTCGATTTTATGATGAGGCTAGCTGTGCCCTCGATTTCAGAAAGCAAGTCAGCT	1533
Qy	501	ValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe	514
Db	1534	GTTCAAGCTGCATCGGAGAGGATACCCACAGGATTTTCATTT	1575

## RESULT 2

```

US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 2021
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1
; LOCATION: (723)..(723)
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

```

Alignment Scores:

Alignment Score:			
Pred. No.	3,978-265	Length:	2021
Score:	2568.50	Matches:	514
Percent Similarity:	95.01%	Conservative:	0
Best Local Similarity:	95.01%	Mismatches:	0
Query Match:	99.09%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-8 (1-514) X US-09-873-409-15 (1-2021)

Qy	1	MettLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	34	ATGATCCTGGGTACTGTCATCACTGTCATATGGAGCGTCCTTCCTTTAATGCCACTG	93
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40

Db 94 GTTTTAGGAGAAATGAGTGATTAACCTTTATTAGTGGGATGCTAGTCCAAACTAACACATAC 153  
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
Db 154 TCTTTCTTCAGGTGACCTGTATATATGTTGGAATAGTGTGCTGCTTGTATTTTGT 213  
Qy 61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
Db 214 TACATACAGATTTCTCTGTGGATTATAACTGCAGCAGCAGACCAAGAGGATTCCAAAA 273  
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100  
Db 274 CAGTTTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTGTATAGCTGTGCATCGGT 333  
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 334 GAACCTTAACACTCGCATGACAGACATTCACAAATCAGTGATGTATTCGAGATAGATT 393  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCGATTGGCTGGCAGTTGGTTGGTGAAG 453  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 454 GGCTGGAAACTCACCTAGTACTCTATCCACGTCCTCTTATATATGCTTCAGCGGCA 513  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 514 GCATGTTCTAGGATGGTCACTCATCTGACCAAGTAAGGAATTAAAGTCCCTATTCGAAGCT 573  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 574 GGCGCTGTGGCAGAGAGCTTGTGCATCAATCCGAACAGTCATAGCCTTTAGGGGCCAG 633  
Qy 201 GluLysGluLeuGln----- 205  
Db 634 GAGAAAGAACTTCAAAGGCTTTCTCTTTAAATATAACAAGATATGCTGGTTTATTTT 693  
Qy 206 -----ArgTyrThrGlnAsnLeuLysAsp 213  
Db 694 CCCAGTGGCTACTAAGTGTGTGTTCTGTTNTTGTAAAGTATACACAGATCTCAAGAT 753  
Qy 214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233  
Db 754 GCAAAGGATTTTGGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGCTGTGTGTAC 813  
Qy 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253  
Db 814 TTCTTTTATGAATGGAACCTATGGACTTGTCTTTTGGTATGGAACCTCTCTGTATCTTAAT 873  
Qy 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273  
Db 874 GGAAACCTGGATATACCATCGGAGCTGTCTTCTGCTGTTTCTTTAGTGTATCCATAGC 933  
Qy 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293  
Db 934 AGTTATTGCAATTCGAGCAGCAGCTCCCTCACTTTTGAAACCTTCGCAATAGCCCGAGAGCT 993  
Qy 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313  
Db 994 GCCTTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATACTTTTCCAGCT 1053  
Qy 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333  
Db 1054 GGATATAAACCTGAATCCATAGAGGAACGTGTGGAATTTTAAATGTTTCTTTCAATAT 1113  
Qy 334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353  
Db 1114 CCATCAAGACCATCTATCAAGATTTCTGAAGGCTCTGAATCTCAGATTAAGTCTGGAGAG 1173  
Qy 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373  
Db 1174 ACAGTGCCTTGGTCTCAATGGCAGTGGGAAGAGTAGTCGGTAGTCCAGCTTCTGCAG 1233

Qy 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393  
Db 1234 AGGTTATATGATCCGGATGATGGCTTTTATCATGGTGGATGAGAAATCAGATCAGAGCTTTA 1293  
Qy 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413  
Db 1294 AATGTGCGGCATTTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG 1353  
Qy 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433  
Db 1354 ACCACCATCAGTAAACATATCAAGTATGGACGAGATGATGTGCTGATGATGAGATGGAG 1413  
Qy 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453  
Db 1414 AGAGCAGCAAGGGAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473  
Qy 454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473  
Db 1474 ACATTTGGTAGGGGNAAGGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533  
Qy 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493  
Db 1534 GCTCGTGCCTTAGTTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTAGCTTCGCCCTG 1593  
Qy 494 AspSerGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSer 513  
Db 1594 GATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACCTCGAAGAGATACCCCCAGGTATTCA 1653  
Qy 514 Phe 514  
Db 1654 TTT 1656  
RESULT 3  
US-09-873-409-14  
; Sequence 14, Application US/09873409  
; Patent No. US20020037522A1  
; GENERAL INFORMATION:  
; APPLICANT: Frank, Markus  
; APPLICANT: Sayegh, Mohamed  
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein  
; FILE REFERENCE: 81994/268611  
; CURRENT APPLICATION NUMBER: US/09/873,409  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 14  
; LENGTH: 3621  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-409-14  
Alignment Scores:  
Pred. No.: 5,61e-263 Length: 3621  
Score: 2552.00 Matches: 507  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.46% Indels: 0  
DB: 10 Gaps: 0  
US-09-873-409-8 (1-514) x US-09-873-409-14 (1-3621)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 34 ATGATCCTGGGTATATCTGGCATCTGCTGGTCAATGGAGCCTGCCTCTTTTAATGCCACTG 93  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 94 GTTTTAGGAGAAATGAGTGAATACCTTTATTAGTGGATGTCTAGTCCAACTAACACATAC 153  
Qy 41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60



D	b	154	TCTTTCTTCAGGTTGGACCCCTGTATTATGTTGGTAATAGGTGTTGCTGCCTCGTCAATTTTTGGT	213
Q	y	61	TyIleGlnIleSerLeuTrpIleileThrAlaAalArgGInThrLyseArgIleAargLys	80
D	b	214	TACATACAGATTTCCTTGTGGATTATAACTGCAGCACACAGACCAGAGGATTCGA AAA	273
Q	y	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly	100
D	b	274	CAGTTTTTTTCATTCAGTTTGCCACAGGACATCGGTGTGTTGTATAGCTGTGTGACATCGGT	333
Q	y	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyVleGlyAspLysIle	120
D	b	334	GAACCTAAACACTCGCATGACAGACATPGACA AAATCAGTGTATGTGGAGATAAGAATT	393
Q	y	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
D	b	394	GCTCTGTTGTTTCAAACAATGTC TACTTTTTGCATTTGGCTGGCAGTTGGTTGGTGAG	453
Q	y	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
D	b	454	GGCTGGAAAACCTACCCCTAGTAGCTATCCACGCTCTCTCTTATAATGGCTTCAGCGCA	513
Q	y	161	AlaCysSerHrgMetValIleSerLeuThrSerLysGluLeuSerAlaTy rSerLysAla	180
D	b	514	GCATGTTCTAGS ATGTGTCATCTCATTTAGCACGATAGGAATTAAGTGGCCTATTTCCAAAAGCT	573
Q	y	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
D	b	574	GGGGCTGTGCCAGAGAGTCTTTGTCATCAATCCGNAACAGTCATAGCCTTTAGGGCCCCAG	633
Q	y	201	GluLysGluLeuGlnArgTy rThrGlnAenLeuLysAspAlaLysAspPheGlyIleLys	220
D	b	634	GAGAAAGAACTTCAAAGGTATACACAGATCTCAAAGATGCCAAAGGATTTTGGCATAAAA	693
Q	y	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTy rPhePheMetAsnGlyTh rTy r	240
D	b	694	AGGACTATAGCTTTCAAAGTGTCCTVTGGTGCTGTGTACTTCTTTATGAATGGAAACCTAT	753
Q	y	241	GlyLeuAlaPheTrpTy rGlyThrSerLeuIleLeuAenGlyGluProGlyTy rThrIle	260
D	b	754	GGACTTGCTTTTTGGTATGGAACCTCCTTGANTCTTATGGNAAACCTGGATATACCATC	813
Q	y	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTy rCysIleGlyAlaAla	280
D	b	814	GGGACTGTTCTTCGCTGTTTCTTTAGTGTAAATCCATAGCAGTATTGTCA TTGGAGCAGCA	873
Q	y	281	ValProHisPheGluThrPheAlalleAlaArgGlyAlaAlaPheHisIlePheGlnVal	300
D	b	874	GTCCCTCAC TTTTGA AACCTTCGCAATAGCCCGAGAGAGCTGCCCTTTCA TATTTTCCAGGTT	933
Q	y	301	IleAspLysLysProSerIleAspAenPheSerThrAlaGlyTy rLvsProGluSerIle	320
D	b	934	ATTGAT AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATTCATTA	993
Q	y	321	GluGlyThrValGluPheLysAenValSerPheAsnTy rProSerArgProSerIleLys	340
D	b	994	GAAGRACTGTGGAAATTA AAAATGTTCTTTCAATTTATTCATCAAGACCATCTATCAAG	1053
Q	y	341	IleLeuLysGlyLeuAenLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
D	b	1054	ATTCTGAAAGGCTCGAATCTCAGAAATTAAGTCTGGAGAGACAGATCGCCTTGGCTCGGCTC	1113
Q	y	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnAArgLeuTy rAspProAspAsp	380
D	b	1114	AATGGCAGTGGGAAGAGTACGTPAGTCCAGCTTCTG CAGAGGTATATGATCCCGAGTAT	1173
Q	y	381	GlyPheIleMetValAspGluAenAspIleArgAlaLeuAsnValArgHisTy rArgAsp	400
D	b	1174	GGCTTTATCATGTGGATGGAATGATCATCAGAGCTTTAAATGTCGGCATATTCGAGAC	1233
Q	y	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAnile	420
D	b	1234	CATATTSAGAGTGTAGTCAAGAGCTGTTTGTGTTGGGACCCACCATCAGTAA CAATATC	1293

Qy	421	LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAla	440
Db	1294	AAGTATATGCGACAGATGATGCTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAT	1355
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	1354	CGGTATGATTTTATCATGGAGTTTCTCTAATAATTTAATACATTTGGTAGGGGAAAAAGGA	1413
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn	480
Db	1414	GCTCAATGAGTGGAGGCGAGAAACAGAGGATCGCAATTTGCTCGTGCCTTAGTTCGAAAC	1473
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	1474	CCCAAGATTCGTATTTTATGATGAGGCTACGTCGTGCCCTGGATTTCAGAAAGCAAGTCAGCT	1533
Qy	501	ValGlnAlaAlaLeuGluLys	507
Db	1534	GTTCAAGCTGCACCTGGAGAAG	1554
RESULT 4			
US-09-873-409-13			
; Sequence 13, Application US/09873409			
; Patent No. US20020037522A1			
; GENERAL INFORMATION:			
; APPLICANT: Frank, Markus			
; APPLICANT: Savedh, A Gened			
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glyc			
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof			
; FILE REFERENCE: 81994/268611			
; CURRENT APPLICATION NUMBER: US/09/873,409			
; CURRENT FILING DATE: 2001-06-05			
; NUMBER OF SEQ ID NOS: 19			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 13			
; LENGTH: 3702			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: No. US20020037522A1e			
; LOCATION: (723)..(723)			
; OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C, G)			

```
Db 274 CAGTTTTTTCATTCACTGTTGGCAGAGGACATCGGCTGGTGGTTGATAGCTGTGCATCGGT 333
Qy 101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 334 GAACTTTAACACATCGCATGACAGACATTTGACAAATCAGTGTGTTATGGAGATAAGATT 393
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTCGATTGGCTGCGCATTGGTTGGTGAAG 453
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 454 GGCCTGAACACTCACCTAGTACTCTATCCAGCTCTCTCTTAATATGGCTTCAGCGCA 513
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaThrSerLysAla 180
Db 514 GCATGTTCTAGGATGCTCATCTCATTTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 574 GGGGCTGTGGCAGAAAGTCTTGTCATCAATCCGAACAGCTCATAGCCTTTAGGGGCCAG 633
Qy 201 GluLysGluLeuGln----- 205
Db 634 GAGAAAGAACTTCAAAGGCTTTTCCTTTTAAATATAACAAGATATGCTGGTTTATTTT 693
Qy 206 -----ArgTyrThrGlnAsnLeuLysAsp 213
Db 694 CCCAGTGGCTACTAAGTTGTGTCTGTTNTTGTAAAGTATACACAGAATCTCAAGAT 753
Qy 214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
Db 754 GCAAAGGATTTTGGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGTGTGTGTAC 813
Qy 234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
Db 814 TTTCTTTATGAATGGAACTTATGGACTTGCTTTTGTATGGAACCTCTTGATTTTAAT 873
Qy 254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
Db 874 GGAAACCTGGATATACCATCGGACTGTCTTCTGTGTTTCTTTAGTGTATCCATAGC 933
Qy 274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
Db 934 AGTTATTGTCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993
Qy 294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
Db 994 GCCTTTTCATATTTTCCAGGTTATTGTATAAGAAACCCAGTATAGATAACTTTTCCAGCT 1053
Qy 314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
Db 1054 GGATATAAACCTGAATCCATAGAAGGAACGTGTGGAAATTTTAAAAATGTTTCTTTCAATTAT 1113
Qy 334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
Db 1114 CCATCAAGACCATCTATCAAGATCTTGAAAGTCTGAACTCTCAGAAATTAAGTCTGGAG 1173
Qy 354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
Db 1174 ACAGTCGCTTGGTGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233
Qy 374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
Db 1234 AGGTTATATGATCCGATGATGGCTTTATCATGTGGATGAGAATGACATCAGAGCTTTA 1293
Qy 394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
Db 1294 AATGTGGGCAATTATCGAGACCATATTTGGATGGTGTAGTCAAGAGCCTGTTTGTTCGGG 1353
Qy 414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
Db 1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGCTGATGAAGAGATGGAG 1413
```

```
Qy 434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
Db 1414 AGAGCAGCAAGGGAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAAATTAAT 1473
Qy 454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
Db 1474 ACATTGGTAGGGGAAAGAGGAGCTCAAAATGAGTGGAGGCGCAAAACAGAGGATCGCAATT 1533
Qy 474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
Db 1534 GCTCGTGCCTTAGTTTCGAAACCCCAAGATTCCTGATTTTAGATGAGGCTACGCTGCCCTG 1593
Qy 494 AspSerGluSerLysSerAlaValGlnAlaLeuGluLys 507
Db 1594 GATTTCAGAAAGCAAGTCAGCTGTTCAAAGTCGACTGGAGAAG 1635
```

## RESULT 5

```
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12
```

```
Alignment Scores:
Pred. No.: 8,09e-172 Length: 3177
Score: 1699.50 Matches: 343
Percent Similarity: 92.70% Conservative: 0
Best Local Similarity: 92.70% Mismatches: 0
Query Match: 65.57% Indels: 27
DB: 10 Gaps: 1
```

US-09-873-409-8 (1-514) x US-09-873-409-12 (1-3177)

```
Qy 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCATCTCATTTGATCCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGGCA 60
Qy 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGTCTTGTTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCCAGGAGAAGACTT 120
Qy 205 Gln----- 205
Db 121 CAAAGGCTTCTCTTTTAAATATAACAAGATATGCTTGGTTTATTTTCCCCAGTGGCTA 180
Qy 206 -----ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 181 CTAAGTTGTGTTCTGTGTTNTTGTAAAGGTATACACAAATCTCAAGATGCAAGAGATTTT 240
Qy 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db 241 GGCAATAAAGAAAGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTACTTCTTTATGAAT 300
Qy 238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
```

Db 301 GGAACCTATGGACTGCTTTTGGTATGGAACTCTCTTGATTTCTTAATGGAGAACCTGGA 360  
Qy TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerTyrCysIle 277  
Db 361 TATACCATCGGAGCTGCTTCTGCTGTTTCTTAGTGTATTCATAGCAGTATTGCAAT 420  
Qy GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297  
Db 421 GGAGCAGCAGTCCCTCACTTGAACCTTCGCAATAGCCCGAGAGCTGCTTTCATATT 480  
Qy PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317  
Db 481 TTCAGGCTTATGATAAGAAACCCAGTATAGATAAATCTTCCACAGCTGGATATAAACCT 540  
Qy 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337  
Db 541 GAATCCATAGAGGAACCTGGGAATTTAAATAATGTTCTTCAATATCATCAAGACCA 600  
Qy 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357  
Db 601 TCTATCAAGATTCTGAAAGCTCTGAATCTCAGAAATTAAGTCTGGAGAGCAGTCGCCCTG 660  
Qy 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377  
Db 661 GTCGGTCTCAATGGCAGTGGGAAGACGATCGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720  
Qy 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397  
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGCGGCAT 780  
Qy 398 TyrArgAspHisIleGlyValValSerGluProValLeuPheGlyThrThrIleSer 417  
Db 781 TATCGACACCATATTGGAGTGGTTAGTCAAGAGCTGTTTGTTCGGGACCAACCATCAGT 840  
Qy 418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437  
Db 841 AACAAATATCAATGATGACGAGATGATGTGACTGATCAAGAGATGGAGAGCAGCAAGG 900  
Qy 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProHisLysPheAsnThrLeuValGly 457  
Db 901 GAGCAAAATCGTATGATTTTATCATGGAGTTTCCATAAATAATTAATCATTTGGTAGGG 960  
Qy 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477  
Db 961 GAAAAAGAGCTCAAATGAGTGGAGGCGCAGAACAGAGGATCGCAATGCTCGTGCCTTA 1020  
Qy 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497  
Db 1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCGCTGGGATTCAGAAAGC 1080  
Qy 498 LysSerAlaValGlnAlaLeuGluLys 507  
Db 1081 AAGTCAGCTGTTCAGCTGCACTGGAGAG 1110

## RESULT 6

US-09-866-866A-5  
; Sequence 5, Application US/09866866A  
; Patent No. US2002010224A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866, 866A  
; PRIOR FILING DATE: 2001-08-30  
; CURRENT APPLICATION NUMBER: 09/584,586  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: PCT/US99/11825  
; PRIOR FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: 60/086,988  
; PRIOR FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.0

; SEQ ID NO 5  
; LENGTH: 4189  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-866-866A-5  
Alignment Scores:  
Pred. No.: 6,47e-160 Length: 4189  
Score: 1590.00 Matches: 307  
Percent Similarity: 76.32% Conservatives: 99  
Best Local Similarity: 57.71% Mismatches: 94  
Query Match: 61.34% Indels: 32  
DB: 10 Gaps: 5

US-09-873-409-8 (1-514) x US-09-866-866A-5 (1-4189)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 148 ATGATTCTGGAACTCTCGCTGCTATTATTCATGGAACATTACTCCCTCTTTGATGCTG 207  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAGCAGAACCCAGTATTCTGCCAAGCAATT 267  
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
Db 268 ACTAATCAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAGAG 327  
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56  
Db 328 ATGCCATATACGCCCTAC-----TATTACACCGGATTTGGTGTGTG 372  
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLys 76  
Db 373 CTGATGTTGCCATCATCCAGGTTTCACTTGGTGGCTGGCAGTCAAGACAGATACAC 432  
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer 96  
Db 433 AAGATTAGGCAGAGATTTTCCATGCTATAATGAATCAGGAGATAGCTGGTGTGTG 492  
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115  
Db 493 CATGATGTTGGGAGCTCAACACCCGCTCACAGATGATGCTCTCCAAATTAATAGCGGA 552  
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
Db 553 ATGGTGACAAAATGGGATGTTTTCAGTCCATACCAACCATTTTACCGCTTGTGCTTATT 612  
Qy 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
Db 613 ATAGGATTTATAGTGGTTGGAGCTAACCTTGTCTATTGCTGCTCAGCCCTCTTATT 672  
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
Db 673 GGATTGTCATCTGCTTTGTGGCAAGGATTTGACTTCATTCTTAATAAGGAACCTCCAG 732  
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluLeuValLeuSerSerIleArgThrValIle 195  
Db 733 GCTTATGCAAAAGCTGGAGCAGTGTCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792  
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215  
Db 793 GCCTTTGGAGGACACAGAGGAGACTTGAAGGTACAAATAAATTTAGAGAGGCTAAA 852  
Qy 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235  
Db 853 AATGTTGGCATAAAGAAAGGATATCACAGCCAGCATTTCCGATAGGCATTGCTACCTGTTG 912  
Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu 255  
Db 913 GTCTATGCATCATATGCATGCGCAATTTCTGTATGGACATCTCTGTGCTCTCAAAATGAA 972  
Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyr 275

```
Db 973 -----TATTCTATTGGAGAGTGGCTACTGCTCTTCTCTATTGTTGGGAGCTTTT 1026
Qy 276 CysileGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db 1027 AGTATTGGACACTTGGCCCAACATAGAGCTTTGCAACGCACAGAGGGGAGCTTTT 1086
Qy 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db 1087 GAAATCTTCAAGATAAATGATAACGAGCCCAAGCAATTCACAGACTTCTCAACAAAGGGCTAC 1146
Qy 316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
Db 1147 AAACACAGACAGTATAATGGGAAACTTAGAGTTTAAATAATGTTCACTCAACTACCCATCG 1206
Qy 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db 1207 AGAAGCGAAGTTCAGATCTTGAGGGCCCTCAATCTGAAGTGAAGCGGACAGCGGTG 1266
Qy 356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db 1267 GCCTTGTTGGCAACAGTGGCTGTGGAAGAACACCAACTGTCCAGCTGATGTCAGAGGCTC 1326
Qy 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db 1327 TAGACCCCCCTGGAGGGCGTGGTCACTATCGACGGACAAAGACATCAATCAATGTG 1386
Qy 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db 1387 AGTATCTGAGGAGATCATTTGTTGGTGTGAGTCAAGAACCTGTGCTGTTGCCACACG 1446
Qy 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
Db 1447 ATGCCCGAGAACATTCGTATGCGCCGAGAGATGTCCACATGATGAGATTGAGAAAGCT 1506
Qy 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db 1507 GTCAAGGAAGCCAAATGCTATGATCTTCATCATGAAACTGCCCCACCAATTTGACACCCCTG 1566
Qy 456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
Db 1567 GTTGGTGAGAGGGCGCGCACTGAGTGGGGGAGAGAAACAGAAATCGCCATGCCCCGG 1626
Qy 476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
Db 1627 GCCTTGGTCCGCAATCCCAAGATCTTTTGTGGAGGAGCCACCTCAGCCCTGGATACA 1686
Qy 496 GluSerLysSerAlaValGlnAlaLeuGluLys 507
Db 1687 GAAAGTGAAGCTGTGGTGCAGCGCGCACTGGATAAG 1722
```

```
RESULT 7
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; PRIOR FILING DATE: 2001-08-30
; CURRENT FILING DATE: 2001-08-30
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-866-866A-1
Alignment Scores:
Pred. No.: 8 59e-159 Length: 3860
Score: 1579.00 Matches: 306
Percent Similarity: 77.04% Conservative: 100
Best Local Similarity: 58.06% Mismatches: 99
Query Match: 60.92% Indels: 22
DB: 10 Gaps: 5
```

```
US-09-873-409-8 (1-514) x US-09-866-866A-1 (1-3860)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGCTGGTGGGAACCTTTGGCTGCCATCATCCATCGGGCTGGAGCTTCCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCCAAATGTCAGGAAATTTAGAGATCTGATGTCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAGAGTGTATCATGATACAGGCTTCTTCATGATCTGGAGGAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCAGATATGCTTATTATTACAGTGAATTTGGTGGGTGCTGGTGGCTGCTTAC 390
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGTGTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATGCGACAGAGATAGGCTGTTTGTATGTCGCCAGATGTTGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGACTTACAGATGATGCTCTAGATTAATAGAGTATTGGTGACAAAT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATCTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACG 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTGGACTGTGAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCTGGCAAGATACTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTGAGAGGCTTGGCAGCAATTTAGAACCTGTGATTGCAATTTGGAGACAA 810
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 811 AGAAAGAACTTGAAAGGTACAAACAAATTTAGAAAGAGCTAAAGAAATTTGGATTAAG 870
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 871 AAAGCTATTACAGCCCAATATTCTATAGTGTGCTTCTCTGCTGATCATGATCTTAT 930
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 931 GCTCTGGCTTCTGTTATGGGACCACTTGGTCTCTCAGGGAA-----TATTCTATT 984
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 985 GGACAAAGTACTCACTGATTTCTTCTGATTAAATTTGGGGCTTTTGTAGTGTGGACAGGCA 1044
```

QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1045 TCTCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1104  
QY 301 IleAspLysLysProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIle 320  
Db 1105 ATTGATTAATGAAGCAAGATTTGACAGCTATTCGAAGAGTGGGCACAAACAGATAATATT 1164  
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1165 AAGGAATTTGGNAATTCAGAAATGTTCACTTCAGTACCCATCTCGAAAGAGTTAAG 1224  
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1225 ATCTTGAAGGCTGAACTGAAGGTGCAGAGTGGGCAGCGTGGCCCTGGTTGGAAC 1284  
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1285 AGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1344  
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1345 GGGATGTGCTAGTGTGATGACAGAGATATTAGGACCAATTAATGTAGGTTTCTACGGAA 1404  
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
Db 1405 ATCATTGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGCCACCCAGATAGTGAACATT 1464  
QY 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1465 CGCTATGGCTGTAATGTGTCACCATGGATGAGATTGAGAAGCTGTCAGAGGAAGCCAA 1524  
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460  
Db 1525 GCCTATGACCTTATCATGAACCTCCCTCATAAATTTGACACCTTGGTTGAGAGAGGG 1584  
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1585 GCCCAGTTGAGTGTGGCAGAACACAGAGGATGCCATTCAGCTGCCCTGGTTCCGAAC 1644  
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1645 CCCAAGATCTCTCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCATG 1704  
QY 501 ValGlnAlaAlaLeuGluLys 507  
Db 1705 GTTCAGGTGGCTCTGGATAAG 1725

## RESULT 8

US-10-044-671-1  
; Sequence 1, Application US/10044671  
; Patent No. US20020177147A1

## GENERAL INFORMATION:

; APPLICANT: Washington State University Research Foundation  
; APPLICANT: Mealey, Katrina  
; APPLICANT: Bentjen, Steven  
; TITLE OF INVENTION: MDRI VARIANTS AND METHODS FOR THEIR USE  
; FILE REFERENCE: 4630-61733  
; CURRENT APPLICATION NUMBER: US/10/044,671  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/261,578  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: US 60/314,829  
; PRIOR FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1

; LENGTH: 4317

; TYPE: DNA

; ORGANISM: Canis familiaris

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (70)..(3912)

## OTHER INFORMATION:

US-10-044-671-1

## Alignment Scores:

Pred. No.: 2,13e-158 Length: 4317  
Score: 1576.00 Matches: 302  
Percent Similarity: 75.61% Conservatives: 98  
Best Local Similarity: 57.09% Mismatches: 105  
Query Match: 60.80% Indels: 24  
DB: 9 Gaps: 4

US-09-873-409-8 (1-514) x US-10-044-671-1 (1-4317)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 217 ATGTTGGTGGGACAATGGCTGCCATCATCCATGGAGTGCACCTCCCTCTCATGATGCTG 276  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
Db 277 GTTTTGGAAACATGACAGATAGCTTTGCAATGTCAGGAATTTCAAGAAACAAACTTTT 336  
QY 35 -----ValGlnThrAsnThrTyrSerPhePheArg----- 44  
Db 337 CCAGTTATAATTAATGAAGTATTACGAACAATACACATTCATCAACCATCTGGAG 396  
QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
Db 397 GAGGAATGACCACGATATGCCATTATTACAGTGGGATCGGTGCTGGCGTGGTGGCT 456  
QY 60 GlyTyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
Db 457 GCTTATCATCCAGGTTTCAATTCGTGCTGGCAGCAGGAGACAGATACTCAAAATTAGA 516  
QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99  
Db 517 AAACAATTTTTCATGCTATCATGCGACAGAGATTTGGCTGGTTTACGTGCATAGCIT 576  
QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
Db 577 GGGGAGCTTAACACCGGCTCACAGACGATGTCTCCAAATCAATGAAGGAATGGCGAC 636  
QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
Db 637 AAAGTTGGAATGTTCTTCAATCAATAGCAACAATTTTTCACCGGTTTATATAGTGGGTT 696  
QY 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
Db 697 ACACCTGGTTGGAAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTCTTGACACTTCA 756  
QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
Db 757 GCGGCCATCTGGGCAAGATATCTATCTTCAATTAAGTAAGAACTCTTGGGCTATGCA 816  
QY 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
Db 817 AAAGCTGGAGCAGTAGCTGAAGAAAGTCTTAGCAGCAATCAGAACTGTGATGCTTTGGA 876  
QY 199 AlaGlnLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
Db 877 GGAACAAAGAAAGAACTTGAAGGTACAAACAAATTTTGAAGAAAGCTAAAGAAATTTGG 936  
QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238  
Db 937 ATAAAGAAAGCTATCAGGCCAACATTTCTATTGGTGGCGCTTCTTATTGATCTATGCA 996  
QY 239 ThrTyrGlyLeuAlaPheTyrPyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258  
Db 997 TCATATGCTCTGGCTTCTGGTATGGGACCTCCTTGGTCTCTCCAGTGAA-----TAT 1050  
QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278  
Db 1051 ACTATTGGACAGGTACTACTGCTCTTCTTCTGTATTAAATTTGGGGCTTTTAGTATTGGA 1110

QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298  
Db 1111 CAGGCATCCCAAGCATTTGAAGCATTTGCAACGCGAGAGGAGCAGCTTATGAATCTTC 1170  
QY 299 GlnValIleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
Db 1171 AGATTAATTGACAAATACCAACCATTCAGCAGCTATTCGAGAGTGACATAACAGAT 1230  
QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
Db 1231 AATATTAAAGGAATTTGGAATTTCAAAATGTTCACTTCAGTTACCTTCTCGAAGAA 1290  
QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358  
Db 1291 GTTAAGATCTTAAAGGCTCTCAACCTGAAGGTTTCAGAGTGGCGAGACAGTGGCGTGT 1350  
QY 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378  
Db 1351 GGAACAGTGGCTGCGGGAAGACGACCGCTGCAGCTGATGCAGAGGCTCTATGACCCC 1410  
QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398  
Db 1411 ACAGATGGCATGCTGCTGATTGATGCACAGGACATTAGGACCAATAATGTAAGGCATCT 1470  
QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418  
Db 1471 CGGGAATTAATCTGCTGCTGAGTCAGGAGCCTGTGTTGTTGCCACCAAGATAGCTGAA 1530  
QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438  
Db 1531 AACATTCGCTATGCGCGCAAAATGTCACCATGGATGAGATGAGAAAGCTGTTAAGGAA 1590  
QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458  
Db 1591 GCCAATGCCATGATTTATCATGAACACTACCTAATAATTTGACACTCTGGTGGAGAG 1650  
QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaLeuVal 478  
Db 1651 AGAGGGGCGCGGTGAGTGGTGACAGAAACAGAGAATCGCCATGCTCGGGCCCTGGTT 1710  
QY 479 ArgAsnProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498  
Db 1711 CGCAACCCCAAGATTTCTGCTGGATGAGGCAACGTCAGCTCGACACTGGAAGTGA 1770  
QY 499 SerAlaValGlnAlaLeuGluLys 507  
Db 1771 GCAGTGGTTCAGGTGGCCCTCGGATNAG 1797

RESULT 9

US-09-917-800A-1424  
; Sequence 1424, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castelle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIORITY APPLICATION NUMBER: US 60/222,040  
; PRIORITY FILING DATE: 2000-07-31  
; PRIORITY APPLICATION NUMBER: US 60/222,880  
; PRIORITY FILING DATE: 2000-11-02  
; PRIORITY APPLICATION NUMBER: US 60/290,029  
; PRIORITY FILING DATE: 2001-05-11  
; PRIORITY APPLICATION NUMBER: US 60/290,645  
; PRIORITY FILING DATE: 2001-05-15  
; PRIORITY APPLICATION NUMBER: US 60/292,336  
; PRIORITY FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1424  
; LENGTH: 4254  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855  
US-09-917-800A-1424  
Alignment Scores:  
Pred. No.: 6,32e-158 Length: 4254  
Score: 1571.50 Matches: 307  
Percent Similarity: 75.95% Conservative: 94  
Best Local Similarity: 58.14% Mismatches: 104  
Query Match: 60.63% Indels: 23  
DB: 10 Gaps: 4  
US-09-873-409-8 (1-514) x US-09-917-800A-1424 (1-4254)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 247 ATGGCTCTGGGAACCTCTCGCTGCTATCATCCACGGAACCCCTGCTTCCCTCTCTGATGCTG 306  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
Db 307 GTGTCGGATACATGACAGATAGTTTACCCCAAGCAGAGACCCGCACTTGTGACCGAGCG 366  
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45  
Db 367 ATTACTAATCAAAAGTGAATCAACAGTACACATACCGTCAGCGACACAGAGTCTTGGAGGAG 426  
QY 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
Db 427 GACATGGCCATGTATGCCCTACTATTACACGGGCAATTTGGTGGCGGTGTGCTCATCGTTGCC 486  
QY 61 TyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80  
Db 487 TACATCCAGGTTTCACTTTGGTGGCTGGCAGTGGGAGACAAATACACAGATTAGGCAG 546  
QY 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100  
Db 547 AAGTTTTTCCATGCCATCATGAATCAGGAGATAGGCTGTTTGACGTGAATGACGCTGGG 606  
QY 101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119  
Db 607 GAGCTCAACACCGCTCTCACAGATGACGCTCCAAATAATTAATGACGGAATTTGGTGCAAA 666  
QY 120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139  
Db 667 CTGGAATGTTCTTTCAGTCCATCAACGACATTTTACCGCGGTTTTATATAGAGTTTATA 726  
QY 140 LysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159  
Db 727 AGTGGTTGGAAGCTAACCTTGTGAATTTTGGCGGTGAGCCCTCTTATTTGGTTGTCTATCT 786  
QY 160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179  
Db 787 GCCATGTGGGCAAGGTACTGCTTCAATTACTTAATAAGGAACCTCCAGGCTTATGCGAAA 846  
QY 180 AlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199  
Db 847 GCTGGAGCAGTTGCCCAAGAAAGTCTTTAGCAGCCATCAGAACTGTGATTTGGTTTGAGGA 906  
QY 200 GlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIle 219

```
Db 907 CAAAAGAGAACTTGAAGAGTCAATAAATAATTAGAAGAGCTAAAGAGTTGGCATA 966
Qy 220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThr 239
Db 967 AAGAAAGCCATCAGCGCAACATTTCCATAGGTATTGCTTACCTACCTGTGGTCTATGGCTCT 1026
Qy 240 TyrGlyLeuAlaPheTptTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThr 259
Db 1027 TATGCACTGGCAATCTGGTATGGACCTCTTGGTCTCTCAATGAA-----TATTCT 1080
Qy 260 IleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAla 279
Db 1081 ATTGGACAAGTGTCTACCGTCTCTCTCTATTATTATTTGGGACCTTTCAGTATTGGACAT 1140
Qy 280 AlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaIlePheGln 299
Db 1141 TTAGCCCAACATAGAAAGCTTTTGCNAATGCAAGAGGGGCGACCTATGAAATCTTCAAG 1200
Qy 300 ValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSer 319
Db 1201 ATAATTGATTAATGAGCCACGATCGACAGCTTCTCAACCAAGGACACAAACCCAGACAGT 1260
Qy 320 IleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIle 339
Db 1261 ATAATGGAAATTTGGAATTTAAATATGTTTACTTCAACTACCCATCACCAGAGTGAAGTT 1320
Qy 340 LysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGly 359
Db 1321 AAGATCTTGAAGGGCCCTCAACCTGAAGCTGAAGAGCGGGCAGACGGTAGCCCTGGTGGC 1380
Qy 360 LeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 379
Db 1381 AACAGTGGCTGTGGGAAGACACAACTGTCCAGCTGTCTGAGAGGCTCTACGACCCCAT 1440
Qy 380 AspGlyPheIleMetValAspGluLysAsnAspIleArgAlaLeuAsnValArgHisTyrArg 399
Db 1441 GAGGCGAGGTCAGTATTGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTCGG 1500
Qy 400 AspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsn 419
Db 1501 GAAATCATTTGGGGTGGTGAAGTCAAGAAACCCGCTGCTGTTGCCACCCAGATTGCCGAAAC 1560
Qy 420 IleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAla 439
Db 1561 ATTCGCTATGCCGAGAAACCTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Qy 440 AsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys 459
Db 1621 AATGCTTATGACTTCATCATGAAACTGCCCCACAAATTTGACACCCCTGGTGGTGAAGA 1680
Qy 460 GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArg 479
Db 1681 GGGCGCAGCTAGTGGGGGACAGAAACAGAGGATCGCCATTGGCCCGGGCCCTGGTCCGC 1740
Qy 480 AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer 499
Db 1741 AACCCCAAGATCTCTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCC 1800
Qy 500 AlaValGlnAlaAlaLeuGluLys 507
Db 1801 GTGGTTCAGGCGCTCTGGATAAG 1824
```

## RESULT 10

US-09-866-866A-3  
; Sequence 3, Application US/09866866A  
; Patent No. US2002010224A1  
; GENERAL INFORMATION:  
; APPLICANT: Sorrentino, Brian  
; APPLICANT: Schuetz, John  
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells  
; FILE REFERENCE: 1340-1-021CIP2  
; CURRENT APPLICATION NUMBER: US/09/866,866A

```
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3
```

Alignment Scores:	7,89e-158	Length:	3860
Score:	1570.00	Matches:	305
Percent Similarity:	76.85%	Conservative:	100
Best Local Similarity:	57.87%	Mismatches:	100
Query Match:	60.57%	Indels:	22
DB:	10	Gaps:	5

US-09-873-409-8 (1-514) x US-09-866-866A-3 (1-3860)

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 211 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGCAAGAAATTTAGAGATCTGATGTCA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 271 AACATCACTAATAAGAGATGATCAATGATATACAGGCTTCTTCATGAATCTGGAGGAAGAC 330
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACACAGATATGCTTATTATACAGTGAATTTGGTGGGTGCTGTTGCTGCTTAC 390
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAAGTTTCATTTTGGTGCCTCGCAGTGGCAGTGGAGACAAATACACAAATTTAGAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATGCGACGAGATAGGCTGTTTGTGTCACCATGTTGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTGGTGACAAAT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGTCTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACCGT 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGTGAAGCTAACCCTTGGATTTGGCCATCAGTCCCTGTTCTTGGACTGTCAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 691 GTCTGGGCAAGATACACTACTTCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGAGCAGTAGCTCAAGAGGCTCTTGGCAGCAATTTAGAACCTGTGATTGCTTGGAGGACAA 810
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 811 AAGAAAGAACTTGAAGGTTACAAACAAATTTTAGAAGAAAGCTAAAGATTGGGATTAAG 870
```



```
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
DB 871 AAAGCTATTACAGCAATATTTCTATAGTGTCTCTCTGCTGATCATCTTAT 930
QY 241 GlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 931 GCTCTGGCTTCTGGTATGGGACCATCTGGTCTCTCAGGGGAA-----TATTCAT 984
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
DB 985 GGACAGTACTCACTGTATCTTCTGTATTAATGGGGCTTTAGTGTGGCAGGCA 1044
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
DB 1045 TCTCCAAGCATTTGAACATTTGCAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 1105 ATTGATAATAAGCCAAGTATTGACAGCTATTGCAAGAGTGGGCACAAACAGATAATATT 1164
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 1165 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTACCCTATCCCATCTCGAAAGATTAG 1224
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1225 ATCTTCAAGGGCTGAACCTGAAGGTGCAGAGTGGGCAGAGCGTGGCCCTGGTTGGAAC 1284
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 1285 AGTGGCTGTGGGAAGAGCAACACAGTCCAGCTGATGCAGAGCGTCTATGACCCACAGAG 1344
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1345 GGATGTCAGTGTGATGCAGCAGGATATTAGACCATTAATGTAAGTTTCTACGGGAA 1404
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
DB 1405 ATCATTTGGTGTGGTGTAGTCAGGAACCTGTATTGTTGCCACCAAGATAGTGAACAATT 1464
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1465 CGCTATGGCCGTGAAATGTCCACCATGTGATGAGATGAGATGAGAAAGCTGTCAAGGA 1524
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
DB 1525 GCCTATGACTTTATCATGAAAACCTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGG 1584
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 1585 GCCCAGTTGAGTGGTGGCAGAGCAGAGGATGCCCATTCGACGTGCCCTGGTTGCAAC 1644
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 1645 CCCAAGATCCTCCTGCTGATGAGGCACCTGACCTTGGACACAGAAAGCGAAGCATG 1704
QY 501 ValGlnAlaAlaLeuGluLys 507
DB 1705 GTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 11

```
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; NAME/KEY: misc feature
; LOCATION: (1)-(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (161)-(677)
; OTHER INFORMATION: 5'-LTR
; NAME/KEY: 5'UTR
; LOCATION: (532)-(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)-(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)-(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)-(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta71m4
```

US-09-306-417-1

```
Alignment Scores:
Pred. No.: 2,68e-157 Length: 8630
Score: 1570.00 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 10 Gaps: 5
```

US-09-873-409-8 (1-514) x US-09-306-417-1 (1-8630)

```
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 1370 ATGGTGGTGGGAACCTTTGGTGCATCCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
DB 1430 GTCTTTGGAGAATGACAGATATCTTTGCCAAATGTCAGGAAATTTAGAAATCTGATGTCA 1489
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 1490 AACATCACTAATAGAGTGTATCATGATGATACAGGCTTCTTCATGAATCTGGAGGAGAC 1549
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 1550 ATGACCAGATATGCTTATTATACAGTGGAAATGGTGGGGTGTGGTTGCTGTCTTAC 1609
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 1610 ATTCAAGTTTCAATTTTGGCTGGCAGCTGGAAGCAAAATACAAAAATTAGAAAACAG 1669
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
DB 1670 TTTTTCATGCTATATATGACAGGAGATAGGCTGTTTGTATGTGCACGATGTGGGGAG 1729
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
```

```
Db 1730 CTTAACACCCGACTTACAGATGATGCTCTTAAGATTAAAGCTTATTGGTGACAAAATT 1789
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 1790 GGAATGTTCTTTAGTCATATGGCAACATTTTCTACGGGTTTATAGTAGGATTTACACGT 1849
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1850 GGTTGGAAGCTAACCCCTTGATTTTGGCCATCAGTCTCTTCTTGAGCTGCTCAGCTGCT 1909
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180
Db 1910 GTCTGGCAAGATATCTATCTTACTATAAGAACTTTAGCGTATGCAAAAGCT 1969
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1970 GGAGCAGTACTGAAGAGGCTTGGCAGCAATTAGAACTGTGATGCTATTTGGAGACAA 2029
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 2030 AAGAAAGACTTGAAGGTACACAAAATTTAGAAGAGCTAAAGAAATTTGGGATAAG 2089
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 2090 AAAGCTATTACAGCCCAATATTCTATAGGTGCTCTTCTCTGCTGATCTATGATCTAT 2149
Qy 241 GlyLeuAlaPheTyrTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 260
Db 2150 GCTCTGGCCCTTGTGATGGGACACCTTGTCTCTCAGGGGAA-----TATTCTATT 2203
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 2204 GGCAAGTACTCACTGATTTCTTCTTCTATTAATGGGGCTTTAGTGTGGACAGGCA 2263
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 2264 TCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 2324 ATTGATTAATGAAGCAAGTATTGACAGCTATTTCGAAGATGGGACACCAACCAAGATAATT 2383
Qy 321 GlyGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 2384 AAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAGAAGTTAAG 2443
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 2444 ATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGGACAGCGGTGGCCCTGTTGGAAC 2503
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380
Db 2504 AGTGGCTGTGGGAAGACACAACTGATCCAGTCCAGTGCAGAGGCTCTATGACCCCAAGAG 2563
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 2564 GGGATGTCAGTGTGTATGGACAGGATATTAGGCCAATAATGTAAAGTTTCTACGGGAA 2623
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 2624 ATCATTTGGTGTGGTGTAGTTCAGGAACCTGTATTGTTGCCACCATGATAGCTGAAACATT 2683
Qy 421 LysTyrGlyArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 2684 CGCTATGGCGGTGAAAATGTCCACCATGGATGAGATTGAGAAAGCTGCAAGGAGGCCAAT 2743
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 2744 GCCTATGACTTTATCATGAAGAACTGCCTCATAAATTTGACACCTTGGTGGAGAGAGGG 2803
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2804 GCCCAGTTTGTGTTGGGCGAAGACAGAGGATCGCCATTGACGTGCGCTGTTGCGAAC 2863
```

```
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2864 CCCAAGATCCTCTGCTGATCAGGCCAGCTCAGCTTGGACACAGAAAGCGAAGCAGTG 2923
```

```
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 2924 GTTCAGGTGGCTCTGGATAAG 2944
```

## RESULT 12

```
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector Sfbeta91mSA1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: (161)..(577)
; LOCATION: (161)..(577)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: msa1 mdlr1 cdna
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2
```

```
Alignment Scores:
Pred. No.: 2,68e-157 Length: 8630
Score: 1570.00 Matches: 305
Percent Similarity: 75.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 10 Gaps: 5
```

```
US-09-873-409-8 (1-514) x US-09-306-417-2 (1-8630)
```

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 1370 ATGGTGGTGGGAACATTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 1429
```

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
 Db 1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAAATGACAGAAATTTAGAAAGATCTGATGCA 1489  
 Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
 Db 1490 AACATCACTAATAGAGAGTATCAATGATACAGGGTCTTTCATGNAATCTGGAGGAGAC 1549  
 Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
 Db 1550 ATGACCCAGGTATGCTTATTATACAGTGCAATTTGGTGCTGGGTGCTGTTGCTGCTTAC 1609  
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
 Db 1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGTGAAGACAAATACACAAATTTAGAAACAG 1669  
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
 Db 1670 TTTTTCATGCTATATATGACAGGAGATAGCTGTTTGTATGTGCACATGTTGGGGAG 1729  
 Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 1730 CTTAACACCCGACTTACAGATGATGCTCTAAAGATTAAATGAAGTTATTGTTGACAAAT 1789  
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1849  
 Qy 141 GlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 1850 GGTGTGAAGCTAACCCCTGTGATTGTTGGCCATCAGTCTCTGTTCTTGAGATGTCAGCTGCT 1909  
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 1910 GTCTGGCAAGACTACTATCTTCTTACTATGATAAGAACTCTTACGATGATGCMAAGCT 1969  
 Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 1970 GGAGCAGTAGCTAGCTGAAGAGCTTTGGCAGCAATTAGAACTGTGATTGCAATTTGGAGGACAA 2029  
 Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLys 220  
 Db 2030 AGAAGAAGACTTGAAGAGGTACAAACAAATTTAGAAAGAGCTTAAAGAAATGGGATAAG 2089  
 Qy 221 ArgThrIleAlaSerLysValSerLeuGlyValAlaValTyrPhePheMetAsnGlyThrTyr 240  
 Db 2090 AAAGCTATTACGCCAATATTCTATAGTGTGCTTCTCTGCTGATCTATGCACTTAT 2149  
 Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
 Db 2150 GCTCTGGCCCTTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCATT 2203  
 Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
 Db 2204 GGACAAGTACTCACTGATTTCTTCTGTATTAATTTGGGGCTTTTAGTGTGGCAGGCA 2263  
 Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaAlaPheHisIlePheGlnVal 300  
 Db 2264 TCTCCAAGCATTTGCAATTTGCAATGCAAGGAGCAGCTTATGCAATCTTCAGATA 2323  
 Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
 Db 2324 ATTGATAAAGCCAAAGTATTGACAGCTATTCGAAGAGTGGGCAACCAACAGATATAAT 2383  
 Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
 Db 2384 AAGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCCATCTCGAAAGAGTTAAG 2443  
 Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
 Db 2444 ATCTGAAGGGCTGAACTGAAGGTGCAAGTGGCAGCAGCGTGGCTTGGTTGGAAC 2503  
 Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380

Db 2504 AGTGCTCTGTGGAAAGAGCACACAGCTGATGAGAGCTCTATGACCCACAGAG 2563  
 Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
 Db 2564 GGGATGGTCAGTGTGATGGACAGATATTAGGACCATAAATGTAAGGTTTCTACGGAA 2623  
 Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
 Db 2624 ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTTGGCACCACGATAGTGAACAACAT 2683  
 Qy 421 LysTyrGlyValAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
 Db 2684 CGCTATGGCCGTAAGATGTACCATGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743  
 Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460  
 Db 2744 GCCTATGACTTATCATGAAACTGCTCATAAATTTGACACCTTGGTGAGAGAGGG 2803  
 Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
 Db 2804 GCCCAGTTGAGTGTGGGAGAGAGGATCGCATTTGCACTGCTGCTGTTGCGCAAC 2863  
 Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
 Db 2864 CCAAGATCCTCTCTGATGAGGCCAGCTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923  
 Qy 501 ValGlnAlaAlaLeuGluLys 507  
 Db 2924 GTTCAGTGGCTCTGGATAAG 2944  
 RESULT 13  
 US-09-769-097-3  
 ; Sequence 3, Application US/09769097  
 ; Patent No. US20020055128A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kimberly Anne Brun  
 ; APPLICANT: Richard James Chenery  
 ; APPLICANT: Harma Ellens  
 ; APPLICANT: John Anthony Feild  
 ; APPLICANT: Lin Yue  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
 ; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF  
 ; FILE REFERENCE: GP-50009-C2  
 ; CURRENT APPLICATION NUMBER: US/09/769,097  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 09/208,809  
 ; PRIOR FILING DATE: 1998-12-09  
 ; PRIOR APPLICATION NUMBER: 09/156,800  
 ; PRIOR FILING DATE: 1999-09-17  
 ; PRIOR APPLICATION NUMBER: US99/20770  
 ; PRIOR FILING DATE: 1999-09-10  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4425  
 ; TYPE: DNA  
 ; ORGANISM: RATTUS RATTUS  
 US-09-769-097-3

Alignment Scores:  
 Pred. No.: 2,3e-157 Length: 4425  
 Score: 1566.50 Matches: 300  
 Percent Similarity: 76.82% Conservative: 101  
 Best Local Similarity: 57.47% Mismatches: 102  
 Query Match: 60.44% Indels: 19  
 DB: 10 Gaps: 4

US-09-873-409-8 (1-514) x US-09-769-097-3 (1-4425)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 499 ATGCTGTGGGAACTCTGGCGGCCATTATCCATGGAAATTCGCTCCCACTTATGATGCTG 558



```
Db 499 ATGCTGCTGGGAACCTCTGGCGCCATTATCCATGGAAATGGCGTCCCATCTTATGATGCTG 558
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db 559 GTCTTTGGGACATGACAGATAGCTTTGCAAT- - - -GTAGGAACAACACCGTAGTAGT 612
Qy 41 SerPhePheArgLeuThr- - - - - - - - - - - - - - - - - - - - - - - - - 46
Db 613 AGTTTCTACAATGCTACAGACATATATGCAAGCTGGAGGAGCAAAATGGCCACGTACGCC 672
Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db 673 TACTATTACAGCGCAATGGTCCGCTGTCTCATCTGCTCCATCATCCAGGTTTCACIT 732
Qy 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db 733 TGGTGCCTGGCAGCTGGGAGACAAATACAAAGATTAGCAGAAAGTTTTTCCATGCCATC 792
Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db 793 ATCAATCAGGAGATAGCTGGTTTGACGTGCATGACGTTGGGAGCTCAACACCCCGCTC 852
Qy 107 Thr- - - - -AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuPheGln 125
Db 853 ACAGATGACGCTCCAAAATTAATGAAGAAATTTGGTGCACAAATTTGGAATGTTCTTTAC 912
Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db 913 GCNATGGCAACATTTTGGTGGTTTATAATAGGATTTACTCGCGCTGGAAAGCTTAAC 972
Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db 973 CTGTGATTTGGCCATCAGCCCTGTCTTGGACTGTCAGCTGTATTTGGGCAAGATA 1032
Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db 1033 TTGCTCTCAATTTACTGATAGGAACCTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTG 1092
Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db 1093 GAAGCTTTAGCAGCCATCAGACTGTGATTCCTTTGGAGGACAAAAGAAAGAACTTGAA 1152
Qy 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db 1153 AGGTACAATAACAATTTGGAAGAAGCTTAAAGGCTTGGGATAAGAAAGCTATACCGGC 1212
Qy 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
Db 1213 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGTCATCATATGCTCTGGCATTTCT 1272
Qy 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db 1273 TATGGGACTTCTTGTGCTCACTCAAAAGAA- - - -TACACTATTGGACNAGTGCTCACT 1326
Qy 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyValAlaAlaValProHisPheGlu 285
Db 1327 GTCTTTTCTGTTAATTTGGAGCATTCAGGTGTGGGAGGAGCATTCACAAATATTGAA 1386
Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db 1387 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTTCAGTATAATGATAAAGCCC 1446
Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 1447 AGTATAGACGCTTCTCAAGAGTGGGACAAACCCGACAAACATACAAAGAAATTTGGAA 1506
Qy 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db 1507 TTCAAAATAATTACATTCACTTCACTCCGCTCTCGAAAGACGCTTCAGATCTTTGAAGGCTC 1566
Qy 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 1567 AACCTGAAGGTGAAGCGGCGCAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAA 1626
```

---

```
Qy 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 1627 AGCAACAACCTGTCCAGCTGTCAGAGGCTTACGACACCCCATAGAGGGCAGGTCACTATC 1686
Qy 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 1687 GACGACAGGACATCAGGACCATCAATGTGAGGTATCTCGGGAAATCAATGGGGTGGTG 1746
Qy 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 1747 ACTCAGAACCGCTGCTGTTTCCACCAACAATTCGCGAAACATTCGCTATGGCCGAGAA 1806
Qy 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 1807 AACGTCACCATCGATGAGTACAGAAAGCTGTCAAGGAAGCAATGCCTATGATTTTCATC 1866
Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyValAlaGlnMetSerGly 465
Db 1867 ATGAAACTGCCCCCAACAATTTGACACCTGTTGGTGTGAGAGAGGGCGCAGCTGAGTGGG 1926
Qy 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 1927 GGACAGAAACAGAGGATGCCATTTGCCGGGCGCTTGGTCCGCAACCCCAAGATCCTTTG 1986
Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu 505
Db 1987 TTGGATGAGGCCACGCTGACCTTGGACACAGAAAGCGAGCGGTGTTCAGGCGGCTCTG 2046
Qy 506 GluLys 507
Db 2047 GATAAG 2052

RESULT 15
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:
Pred. No.: 6,64e-157 Length: 4643
Score: 1562.50 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 99
Query Match: 60.28% Indels: 23
DB: 9 Gaps: 6

US-09-873-409-8 (1-514) x US-10-072-621-2 (1-4643)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGTGTGTGGGAACCTTGTGCTGCCATCATCTCCTGCGGCTGGACTTCTCTCATGATGCTG 634
Qy 21 ValLeuGlyGluMetSerAsp- - - - - - - - - - - - - - - - - - - - - - - - - 31
Db 635 GTGTTTGGAAATGACATATCTTTGCAATGCAAGAAATTTAGAAATCTGATGTCA 694
Qy 32 GlyCysLeuValGlnThr- - - - - - - - - - - - - - - - - - - - - - - - - 43
```

```
Db 695 AACATCACTAATAGAGTGTATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC 754
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAGGTATGCCATTATTATACAGTGAATTTGGTCTGGGGTGTGGTCTGTCTTAC 814
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTTCAGGTTTCATTTTGGTCCCTGGCAGCTGGAGACAAATACACAAAATTAGAAAACAG 874
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATGCACAGGAGATAGGCTGGTTTCATGTCACGATGTTGGGAG 934
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACACCCGACTTACAGATGATGCTCCCAAGATTAAATGAAGGAATTTGGTGACAAAAT 994
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTTCAGTCATGGAACCAATTTTTCAGTGGGTTTATAGTAGGATTACACGT 1054
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTGGAAGCTAACCCCTTGATTTTGGCCATCAGTCTCTGTTGGAGCTGTCAGCTGCT 1114
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGlnLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACCTATCTTCAATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174
Qy 181 GlyValAlaAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTACTGAAGAGGCTTGGCAGCAATTAGAACTGTGATTCATTTGGAGGACAA 1234
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 1235 AAGAAAGAACTTGAAGGTACACAAAAATTTAGAAAGCTAAAGAAATTTGGATAAAG 1294
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGlyThrTyr 240
Db 1295 AAAGCTATTACAGCCCAATTTCTATAGGTGCTGCTTTTCTGCTGATCTATGCACTCTAT 1354
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1355 GCTCTGGCCCTTCTGGTATGGACCACTTGTCTCTCAGGGGAA-----TATTCATTT 1408
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1409 GGACAAGTACTCACTGTA---TTTCTGTATTAATTTGGGGCTTTTAGTGTGGACAGGCA 1465
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1466 TCTCCAAAGCATTTGAAGCATTTGCAAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1525
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1526 ATTGATAATAGCCCAAGATTTGACGCTATTGCAAGGTGGGACCAACCCAGATATATT 1585
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1586 AAGGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAGAAAGTTAAG 1645
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1646 ATCTTGAAGGCTCTGAACCTCAAGCTGACAGTGGGACAGCGTGGCCCTGGTTGGAAC 1705
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1706 AGTGGCTGTGGGAAGACACACAGTCCAGCTGATGAGAGGCTCTATGACCCACAGAG 1765
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
```

```
Db 1766 GGGATGGTCAGTGTTCATGATGACAGCAGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1825
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1826 ATCAATTGGTGTGGTGAAGCAGCACTGTATTGTTTCCACCACGATAGCTGAAAACATT 1885
Qy 421 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1886 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGCAAGGAAGCCAAT 1945
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 1946 GCCTATGACTTTTATCATGAAACTGCCCTCATAAATTTGACACCCCTGTTGGAGAGAGGG 2005
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2006 GCCAGTTGAGTGGTGGGAGAGCAGAGATGCGCATTTGCAGCTGCCCTGGTTCCGAAC 2065
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2066 CCCAAGATCCTCTGCTGCTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2125
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 2126 GTTCAGGTGGCTCTGGATAAG 2146
```

Search completed: March 31, 2003, 16:15:25  
Job time : 246.092 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:24:25 ; Search time 6948.72 Seconds  
(without alignments)  
2152.747 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592

Sequence: 1 MILGTLASLVNGACILPLMLP.....SESKSAVQAALKEKTPRYSF 514

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003.115418.19228/app\_query.fasta\_1.7544  
-DB=GenEmbl -OFT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 54887 @runat\_27032003.115418.19228 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

RESULT 1

ALIGNMENTS

Result No.	Score	Query Match %	Length	DB ID	Description
1	2592	100.0	1940	6	AX339034 Sequence
2	2568.5	99.1	2021	6	AX339033 Sequence
3	2552	98.5	3621	6	AX339032 Sequence
4	2528.5	97.6	3702	6	AX339031 Sequence
5	2460.5	94.9	3699	6	AX478104 Sequence
6	1699.5	65.6	3177	6	AX339030 Sequence
7	1590	61.3	4189	6	AX322791 Sequence
8	1590	61.3	4298	10	MUSMDR
9	1585	61.1	4279	6	AX105082 Sequence
10	1581	61.0	4279	6	AX105078 Sequence
11	1581	61.0	4279	6	AX105080 Sequence
12	1580.5	61.0	4195	6	AX108656 Sequence
13	1560	61.0	4045	12	AF269224
14	1579	60.9	3860	6	AX322787 Sequence
15	1579	60.9	4378	6	E02326
16	1579	60.9	4646	6	AX391099 Sequence
17	1578.5	60.9	4280	10	CRUPGPII
18	1576	60.8	3888	10	AY082609
19	1576	60.8	4233	6	AR123273 Sequence
20	1576	60.8	4279	6	AX105057 Sequence
21	1576	60.8	4317	4	AF045016
22	1576	60.8	4317	6	AX105059 Sequence
23	1576	60.8	4669	6	AR091275 Sequence
24	1576	60.8	4669	6	AR203322 Sequence
25	1572	60.6	4186	6	AX108654 Sequence
26	1571.5	60.6	4254	6	AX401748 Sequence
27	1571.5	60.6	4927	10	RATMDRM
28	1571.5	60.6	4927	10	AF257746
29	1570.5	60.6	4323	10	AF286167
30	1570	60.6	3860	6	AX322789 Sequence
31	1570	60.6	4646	6	AX336420 Sequence
32	1570	60.6	4646	6	AX336708 Sequence
33	1570	60.6	4646	6	I49610
34	1570	60.6	4646	9	HUMMDR1
35	1570	60.6	4669	6	I08557
36	1570	60.6	6505	6	AR028671 Sequence
37	1570	60.6	8630	6	AX012320 Sequence
38	1570	60.6	8630	6	AX012321 Sequence
39	1570	60.6	9318	6	AR028672 Sequence
40	1568.5	60.5	4264	6	AR051647
41	1568.5	60.5	4264	6	AR051650 Sequence
42	1565.5	60.4	4356	10	MUSMDRAA
43	1565	60.4	4296	10	CRUPGPI
44	1564	60.3	4304	10	CRUPGPI165
45	1563.5	60.3	3489	4	AB029153 Felis cat

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES



AX339034  
LOCUS AX339034 1940 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 16 from Patent WO0194400.  
ACCESSION AX339034  
VERSION AX339034.1 GI:18129126  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
homologue on chromosome 7p15-21 and uses thereof  
JOURNAL Patent: WO 0194400-A 16 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)  
FEATURES  
Location/Qualifiers  
source 1. 1940  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 530 a 394 c 448 g 568 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.68e-192 Length: 1940  
Score: 2592.00 Matches: 514  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-873-409-8 (1-514) x AX339034 (1-1940)  
Qy 1 MetileuGlyIleLeuAlaSerleuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 34 ATGATCTGGGTACTGGGCATCACTGGTCAATGGAGCGCTTCCTTCTTAATGCCACTG 93  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 94 GTTTTAGGAGAAATGAGTGATTAATCTTATAGTGAGTGTCTAGTCCAACTAACACATAC 153  
Qy 41 SerPheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60  
Db 154 TCTTTCTTCAGGTTGACCTGTATTATGTTGGAATAGGTGTGCTGCTTGGTATTTGGT 213  
Qy 61 TyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrIleArgIleArgIys 80  
Db 214 TACATACAGATTTCCTTGTGGATTATTAATCTCAGCAGCAGACCAAGAGGATTTCGAAA 273  
Qy 81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly 100  
Db 274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGTAGCTGTGACATCGGT 333  
Qy 101 GluLeuAsnThrArgMetThrAspIleAspIysIleSerAspGlyIleGlyAspIle 120  
Db 334 GAACCTTAACACTCGCATGACAGACATTGACAAATCAGTGATGTTATTTGGAGATAAGATT 393  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValIys 140  
Db 394 GCTCTGTTGTTTCAAAACATGCTACTTTTTTCGATTGGCCCTGGCAGTTGGTGGTGAAG 453  
Qy 141 GlyTyrIysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 454 GGCTGGAAACTCACCTAGTGACTCTATCCACGCTCTCTCTTATTAATGGCTTCAGCGGCA 513  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerIysGluLeuSerAlaTyrSerIysAla 180  
Db 514 GCATGTTCTAGGATGGTTCATCTCATTCACCAAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573  
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 574 GGGGCTGTGGCAGAGAGTCTTGTTCATCAATCCGAACAGTCATAGCCTTTTAGGGCCCG 633

Qy 201 GluIysGluLeuGlnArgTyrThrGlnAsnLeuIysAspAlaIysAspPheGlyIleLys 220  
Db 634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAGAGGATTTTGCATAAAA 693  
Qy 221 ArgThrIleAlaSerIysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 694 AGGACTATAGCTTCAAAGGTCTCTTGGTGTGTGTCTTTATGAATGAACCTAT 753  
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
Db 754 GGACTTGCTTTTGGTATGGAACCTCTTGATTTCTTAATGGAGAACCTGGATATACCATC 813  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 814 GGGACTGTTCTTCTGCTGTTCTTTCTTTAGTGAATCCATGACGATTTATTCATTGGAGCAGCA 873  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 874 GTCCCTCACCTTTGAAACCTTCGCAATAGCCGAGGAGCTGCCCTTTCATATTTTCCAGGTT 933  
Qy 301 IleAspIysIysProSerIleAspAsnPheSerThrAlaGlyTyrIysProGluSerIle 320  
Db 934 ATTGATAAGAAACCCAGTATAGATAAATCTTCCACAGCTGGATATAACCTGAATCCATA 993  
Qy 321 GluGlyThrValGluPheIysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 994 GAAGGAACCTGTGGAATTTAAATAATGTTCTTTCATATTCATCAAGACCATCTATCAAG 1053  
Qy 341 IleLeuIysGlyLeuAsnLeuArgIleIysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1054 ATTCTGAAGGTCTGAATCTCAGAATTAAGCTGGAGAGACAGTCGCTTGGTGGGTCTC 1113  
Qy 361 AsnGlySerGlyIysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1114 AATGGCAGTGGGAGAGTACGCTAGTCCAGCTTCTCGAGAGGTTATATGATCCGATGAT 1173  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1174 GGCTTTTATCATGCTGGATGAGATGACATCAGAGCTTTTAATGTGGGCATTTATCGAGAC 1233  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
Db 1234 CATATTGGAGTGGTTAGTCAAGAGCCCTGTTTCTCGGGACCCACCATCAGTAACAATATC 1293  
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAGGAGCAAT 1353  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnIysPheAsnThrLeuValGlyGlyLysGly 460  
Db 1354 GCGTATGATTTTATCATGAGATTTCTCTAATAATTTAATAATTTAGTGGGAAAAAGGA 1413  
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1414 GCTCAATAGTGGAGGGCAGAAACAGAGATCGCAATTCCTGCTGCTTAGTTTCGAAC 1473  
Qy 481 ProIysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1474 CCCAAGATTTCTGATTTTAGATGAGGCTACGCTGCCCCCTGGATTCCAGAAAGCAAGTCAGCT 1533  
Qy 501 ValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe 514  
Db 1534 GTTCAAGCTGCATGGAGAGGATACCCCCAGGTATTCATTT 1575

RESULT 2  
AX339033  
LOCUS AX339033 2021 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 15 from Patent WO0194400.  
ACCESSION AX339033  
VERSION AX339033.1 GI:18129125  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Frank, M.H. and Sayegh, M.H.  
TITLE A gene encoding a multidrug resistance human p-glycoprotein  
JOURNAL homologue on chromosome 7p15-21 and uses thereof  
Patent: WO 0194400-A 15 13-DEC-2001;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

FEATURES

source Location/Qualifiers  
1..2021  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 547 a 406 c 462 g 605 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1..89e-190 Length: 2021  
Score: 2568.50 Matches: 514  
Percent Similarity: 95.01% Conservative: 0  
Best Local Similarity: 95.01% Mismatches: 0  
Query Match: 99.09% Indels: 27  
DB: Gaps: 1

US-09-873-409-8 (1-514) x AX339033 (1-2021)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	34	ATGATCTGGGTATACCTGGCATCACTGGTCAATGGAGCGCTTCCTTTAAATGCCACTG	93
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	94	GTTTTAGGAGAAATGAGTGATACCTATTATTAGTGGATGCTAGTCCAAACTAACACATC	153
Qy	41	SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePheGly	60
Db	154	TCCTTCTCAGGTGACCTGATATTATTGTAATAGTGGTGTGCTGCTTGAATTTTGGT	213
Qy	61	TyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrIleArgIleArgLys	80
Db	214	TACATACAGATTTCCTGTGGATTATAAATGCAGCAGCAGCAGCAAGAGGATTCGAAAA	273
Qy	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGly	100
Db	274	CAGTTTTTTCATTCAGTTTTGGCAGCAGCATCGGCTGGTTGATAGCTGTGATCGGT	333
Qy	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	334	GAACCTTAACACTCGCATCAGACATTCGACAAATCAGTCATGATGGATTTGGAGATTAAGATT	393
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	394	GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGTTGGTGAAG	453
Qy	141	GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	454	GGCTGGAAACTCACCTAGTAGTACTCTATCCACGCTCTCCTCTTAATATGGCTTCAGCGCA	513
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	514	GCATGTTCTAGGATGGTCATCTCATGGACAGTAAGGAATTAAGTGCCCTATTCCAAGCT	573
Qy	181	GlyValaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	574	GGGGCTGTGCAGAGAAGCTCTTGTCATCAATCCGAACAGTCATAGCTTTTAGGCCCCAG	633
Qy	201	GluIysGluLeuGln-----	205
Db	634	GAGAAAGAACTTCAAAAGGCTCTTTCCTTTTAAATATAACAAGATATGCTGGTTTATTTT	693
Qy	206	-----ArgTyrThrGlnAsnLeuLysAsp	213
Db	694	CCCCAGTGGCTACTAAGTTGTTCTGTTTWTGTTAAGGTATACACAGAAATCTCAAGAT	753

Qy	214	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	233
Db	754	GCAAGAGATTTTGGCATAAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGAC	813
Qy	234	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsn	253
Db	814	TTCTTTATGATGGAACCTATGGACTTGTCTTTTGGTATGGAACCTCCTTGATCTTAAAT	873
Qy	254	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	273
Db	874	GGAGAACTGGATATACCATCGGACTGTTCTTGTCTTCTTTAGTGTAAATCCATAGC	933
Qy	274	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	293
Db	934	AGTTATTGCAATGGAGCAGCAGCTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCT	993
Qy	294	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	313
Db	994	GCCTTTTCATATTTTCCAGGTTATTGATAGAAACCCAGTATAGATAACTTTTCCAGCT	1053
Qy	314	GlyTyrLysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyr	333
Db	1054	GGATATAAACCTGAAATCCATAGAGGAACCTGTGGAATTTTAAAAATGTTCTTTCAATTAT	1113
Qy	334	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	353
Db	1114	CCATCAAGACCATCTATCAAGATTTCTGAAAGGTCTGAAATCTCAGAATTAAGTCGAGAG	1173
Qy	354	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	373
Db	1174	ACAGTCGCCTTGTGCGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG	1233
Qy	374	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	393
Db	1234	AGTTATATGATCCGATCGATGCTTATCATGGTGGATGAGAAATCAGATCAGAGCTTTA	1293
Qy	394	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	413
Db	1294	AATGTGCGCATTTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCGCTGTTTGTTCGGG	1353
Qy	414	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu	433
Db	1354	ACCACCATCAGTAACAAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG	1413
Qy	434	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	453
Db	1414	AGAGCAGCAGGAGGAGCAATGCGTATGATTTTATCATGGAGTTTCTTAATAATTTAAT	1473
Qy	454	ThrLeuValGlyLysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIle	473
Db	1474	ACATTGTTAGGGGAAAAAGGAGCTCAAAATGAGTGGAGGGGAGAGGATCGCAATT	1533
Qy	474	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	493
Db	1534	GCTCGTGGCTTAGTTCGAAACCCCAAGATTTGATTTTATAGTGGGCTACGCTCGCCCTG	1593
Qy	494	AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer	513
Db	1594	GATTTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAGGATACCCCGAGTATTCA	1653
Qy	514	Phe 514	
Db	1654	TTT 1656	

RESULT 3  
AX339032  
LOCUS  
DEFINITION Sequence 14 from Patent WO0194400.  
ACCESSION AX339032  
VERSION AX339032.1 GI:18129124  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens



FEATURES		Location/Qualifiers	
source	1..3702	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT	1098 a 708 c 856 g 1039 t	1 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	5.1e-187	Length:	3702
Score:	2528.50	Matches:	507
Percent Similarity:	94.94%	Conservative:	0
Best Local Similarity:	94.94%	Mismatches:	0
Query Match:	97.55%	Indels:	27
DB:	6	Gaps:	1
US-09-873-409-8 (1-514) x AX339031 (1-3702)			
Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProIleuMetProLeu	20
Db	34	ATGATCCTGGGTATACCTGGCATCACTGGTCAATGGAGCCCTGCCTTCCTTTAAATGCCACTG	93
Qy	21	ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr	40
Db	94	GTTTTAGGAGAAATGAGTGATTAACCTTATTAGTGGATGTCCTAGTCCAACTAACACATAC	153
Qy	41	SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly	60
Db	154	TCITTTCTCAGGTGACCCCTGTATTATGTTGGAATAGGTGCTGCTGCCITTGATTTTGGT	213
Qy	61	TyrIleGlnIleSerLeuThrIleIleThrAlaAlaArgGlnThrIlysArgIleArgLys	80
Db	214	TACATACAGATTTCCCTGTGGATTTAATCTGCAGCACAGACAGCAAGAGGATTCGAAA	273
Qy	81	GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly	100
Db	274	CAGTTTTTTCATTCAGTTTGGCACAGGACATCGCGCTGGTTGATAGCTGTGACATCGGT	333
Qy	101	GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	334	GAACCTTAACACTCGCATGACAGACATTCACAAATCAGTGTGTTGGAGATAAGATT	393
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	394	GCTCTGTGTTTCAAAACATGTCTACTTTTTTCGATTTGGCCTGGCAGTTGGTGAAG	453
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProIleuIleMetAlaSerAlaAla	160
Db	454	GGCTGGAACCTCACCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA	513
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	514	GCATGTTCTAGGATGGTCACTCATTCAGCCAGTAAGGAATTAAGTGCCTATTCCAAAGCT	573
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	574	GGGGCTGTGGCAGAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCAG	633
Qy	201	GluLysGluLeuGln-----	205
Db	634	GAGAAAGAACTTCAAAGGCTCTTCCCTTTTAAATATAACAAGATATGCTGGTTTATTTT	693
Qy	206	-----ArgTyrThrGlnAsnLeuLysAsp	213
Db	694	CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTGTAAAGGTATACAGAAATCTCAAAGAT	753
Qy	214	AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr	233
Db	754	GCAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGTGTGTAC	813
Qy	234	PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn	253
Db	814	TTCTTTATGAATGGAACCTATGGACTTGCTTTTGGTATGGAACCTCCTTGATTTCTTAAT	873

Qy	254	GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer	273
Db	874	GGAGAACCTGGATATACCATCGGAGCTGTTCTTGTGTTCTTTAGTCTAATCCATAGC	933
Qy	274	SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla	293
Db	934	AGTTATTGCAATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCGAGAGCT	993
Qy	294	AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla	313
Db	994	GCCTTTTCATATTTCCAGGTTATTGATAAGAAACCAGATATAGATAAATTTTCCACAGCT	1053
Qy	314	GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr	333
Db	1054	GGATATAAACCTGAATCCATAGAAGGAACCTGTGGAATTTAAAAATGTTCTTTCAATTAT	1113
Qy	334	ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu	353
Db	1114	CCATCAAGACCATCTATCAAGATTCTGAAGGTCTGAATCTCAGAATTAACTCTGGAGAG	1173
Qy	354	ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln	373
Db	1174	ACAGTCGCCCTTGGTCGGTCTCAATGGCAGTGGGAAGATACGCTAGTCCAGCTTCTGCAG	1233
Qy	374	ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu	393
Db	1234	AGGTTATATGATCCGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA	1293
Qy	394	AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly	413
Db	1294	AATGTGCGGCATTTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTGTTCGGG	1353
Qy	414	ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu	433
Db	1354	ACCACCATCAGTACACATATCAAGTATGACGAGATGATGCTGACTGATGAGAGATGGAG	1413
Qy	434	ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn	453
Db	1414	AGAGCAGCAAGGGAAGCAAAATCGTATGATTTTATCATGGAGTTTCTTAATAAATTTAAT	1473
Qy	454	ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle	473
Db	1474	ACATTTGGTAGGGGAAAGAGCTCAAATGAGTGGAGGCGCAAAACAGAGGATTCGCAAT	1533
Qy	474	AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu	493
Db	1534	GTCGTGCTTGTAGTTCGAACCCCAAGATTCGATTTTAGATGAGGCTAGCTGCGCTG	1593
Qy	494	AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys	507
Db	1594	GATTCAAGAAAGCAAGTCAGCTGTTCAAGCTGCACCTGGAGAAG	1635
RESULT 5			
AX478104	AX478104	3699 bp	DNA linear PAT 12-AUG
LOCUS	Sequence 34 from Patent WO0240541.		
DEFINITION	AX478104		
ACCESSION	AX478104.1		
VERSION	GI:22217064		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Tang, Y.T., Yue, H., Nguyen, D.B., Hafalia, A.J., Elliott, V.S., Lu, Y., Walla, N.K., Yao, M.G., Baughn, M.R., Gandhi, A.R., Ding, L., Sanjanwala, M., Ramkumar, J., Arvizu, C., Gietzen, K.J., Lal, P.G., Azimzai, Y., Khan, F.A., Thangavelu, K., Thornton, M., Lu, D.A., Tribouley, C.M., Warren, B.A., Ison, C.H., Das, D., Raumann, B.E., Policky, J.L. and Kearney, L.		
Transposers and ion channels			
Patent: WO 0240541-A 34 23-MAY-2002;			
Incyte Genomics, Inc. (US)			
TITLE			
JOURNAL			

## FEATURES

Location/Qualifiers

source

1. .3699  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /notes="Incyte ID No: 7472030CB1"

BASE COUNT 1116 a 707 c 860 g 1016 t

## ORIGIN

## Alignment Scores:

Pred. No.: 9,79e-182 Length: 3699  
 Score: 2460.50 Matches: 495  
 Percent Similarity: 95.40% Conservatives: 3  
 Best Local Similarity: 94.83% Mismatches: 9  
 Query Match: 94.93% Indels: 15  
 DB: 6 Gaps: 2

US-09-873-409-8 (1-514) x AX478104 (1-3699)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 145 ATGATCTGGGTACTGACATCACTGTCAATGGAGCCTGCCTTCTTAAATGCCACTG 204  
 Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
 Db 205 TGTATAGGAGAAATGAGTGAATACCTTATTAGTGGATGCTAGTCCACACTAACACACA 264  
 Qy 41 SerPhePheArg-----LeuThr 46  
 Db 265 AATTATCAGAACTGACTCAGCTCAAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324  
 Qy 47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
 Db 325 CTGTATTATGTTGAATAGGTGCTGCTGCTTGTATTTGGTTACATACAGATTCTCTTG 384  
 Qy 67 TrpIleLeuThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86  
 Db 385 TGGATTATACTGCAGCAGCAGACACCAAGAGGATTCGAAACAGTITTTTCATTCAGTT 444  
 Qy 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
 Db 445 TTGGCAGAGGACATCGGCTGGTGTGTAGCTGTGACATCGGTGAACCTTACACTCGCATG 504  
 Qy 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
 Db 505 ACAGATGACATTGACAAATCAGTGTGTTGAGATGAGATAAGATTGCTCTGTTGTTCAA 564  
 Qy 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145  
 Db 565 AACATGTCATCTTTTCGATTGGCCTGGCAGTTGGTTGGTGAAGGGCTGGAAACTCACC 624  
 Qy 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165  
 Db 625 CTAGTGACTCTATCCAGCTCTCTCTTATATGCTTCAGCGGCAGCATGTTCTAGGATG 684  
 Qy 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
 Db 585 GTCACTCATTCAGCAGTAAGAAATTAAGTCCCTATTCCAAAGCTGGGGCTGGCGAGAA 744  
 Qy 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205  
 Db 745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCGCAGGAGAGAACTTCAA 804  
 Qy 206 ArgTyrThrGlnAsnLeuLysAspAlaLysPheGlyIleLysArgThrIleAlaSer 225  
 Db 805 AGGTATACACAGAAATCTCAAGATGCAAGGATTTTGGCATATAAAGAGACTATAGCTTCA 864  
 Qy 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245  
 Db 865 AAGTGTCTCTGTGCTGCTGCTGCTCTTTATGAATGAAGAACTATGACTGCTTTTGG 924  
 Qy 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265  
 Db 925 TATGGAACTCTCTTGATTCTTTAATGGAGAACTCGATATACCATCGGAGCTGTTCTTGCT 984

Qy 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
 Db 985 GTTTTCTTTAGTGTATCCATAGTAGTATTATTCATTGGAGCAGCAGTCCCTCACTTTGAA 1044  
 Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305  
 Db 1045 ACCTTCGCAATAGCCGAGAGCTGCTTTTATTTTCCAGGTATTATGATAGAACC 1104  
 Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGlyThrValGlu 325  
 Db 1105 AGTATAGGTAACTTTTCCAGCTGGATATAACCTGAATCCATAGAGGAACCTGGGAA 1164  
 Qy 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345  
 Db 1165 TTTTAAATGTTTCTTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAGGCTCTG 1224  
 Qy 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365  
 Db 1225 AATCTCGGAATTAAGTCTGGAGAGACAGTCGCTTGTGCTGCTCAATGGCAGTGGGAAG 1284  
 Qy 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385  
 Db 1285 AGTACGGTAGTCCAGCTCTTCGACAGGTTATATGATCCGATGATGGCTTTATCATGGTG 1344  
 Qy 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405  
 Db 1345 GATGAGAATGACATCAGAGCTTTAATGTGGGCATTATCGAGACCATATGGAGTGGTT 1404  
 Qy 406 SerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 425  
 Db 1405 AGTCAAGAGCCTGTTTGTTCGGACACCATCAGTAACAATATCAAGTATGACGAGAT 1464  
 Qy 426 AspValThrAspGluGluMetGluArgAlaAlaAatGluAlaAsnAlaTyrAspPheIle 445  
 Db 1465 GATGTCACTGATGAAGATGGAGAGCAGCAGGAGGAGCAAAATCGGTATGATTATTC 1524  
 Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465  
 Db 1525 ATGAGATTTCTTAATAATTAATATCATTTGGTAGGGGAAAAGGAGCTCAATGAGTGA 1584  
 Qy 466 GlyGlnGlnArgIleAlaAlaAlaArgAlaLeuValArgAsnProLysIleLeuIle 485  
 Db 1585 GGGCAGAAAACAGAGGATCGCAATTGCTGCTGCTTATGTTAGTTCGAAACCCCAAGATTCTGATT 1644  
 Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505  
 Db 1645 TTAGTAGGAGCTACGCTGCTCCCTGGATTGAGAAAGCAAGTCACTGTTCAAGCTGCACTG 1704  
 Qy 506 GluLys 507  
 Db 1705 GAGAAG 1710

## RESULT 6

LOCUS AX339030 3177 bp DNA linear PAT 09-JAN-2002  
 DEFINITION Sequence 12 from Patent WO0194400.  
 ACCESSION AX339030  
 VERSION AX339030.1 GI:18129122  
 KEYWORDS human.  
 SOURCE Homo sapiens

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 Frank, M.H. and Sayegh, M.H.

A gene encoding a multidrug resistance human p-glycoprotein

homologue on chromosome 7p15-21 and uses thereof

Patent: WO 0194400-A 12 13-DEC-2001;

THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)

## FEATURES

Location/Qualifiers  
 1. .3177  
 /organism="Homo sapiens"

BASE COUNT 970 a 601 c 735 g 870 t 1 others  
ORIGIN /db\_xref="taxon:9606"

Alignment Scores:  
Pred. No.: 11e-122 Length: 3177  
Score: 1699.50 Matches: 343  
Percent Similarity: 92.70% Conservative: 0  
Best Local Similarity: 92.70% Mismatches: 0  
Query Match: 65.57% Indels: 27  
DB: 6 Gaps: 1

US-09-873-409-8 (1-514) x AX339030 (1-3177)

```
QY 165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
Db 1 ATGGTCATCTCATTGACCACTAGGAATTAAAGTCCCTATTCCAAAGCTGGGGCTGGCA 60

QY 185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
Db 61 GAAGAAGCTTGTGCATCAATCCGAACAGTCATAGCCTTTAGGGCCCGAGAGAAAGAACTT 120

QY 205 Gln----- 205
Db 121 CAAGGTCCTTCCTTTAAATATATAACAAGATATGCTTGCTGGTTTATTTCCCCAGTGGCTA 180

QY 206 -----ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db 181 CTAAGTTGTGTTCTGTTNTTGTAAAGTATACACAGAACTCTCAAGATGCAAGGATTTT 240

QY 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsn 237
Db 241 GGCATAAAAGGACTATAGCTTCAAAAGTGTCTCTGGTGTGTACTTCTTTATGAAT 300

QY 238 GlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAnGlyGluProGly 257
Db 301 GGAACCTATGGACTTGCTTTTGGTATGGAACCTCTCTGTGATCTTAAATGGAGAACCTGGA 360

QY 258 TyrThrIleGlyThrValLeuAlaValPheSerValIleHisSerTyrCysIle 277
Db 361 TATACCATCGGACTGTTCTTGCTGTTTTCTTAGTGTATTCATCAGCAGTATTGCATT 420

QY 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db 421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCCGAGGAGTGCCTTTTCATATT 480

QY 298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db 481 TTCAGGTTATTGATAAGAAACCCAGTATAGATAAATTTTCCACAGCTGGATATAAACCT 540

QY 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerAtgPro 337
Db 541 GAATCCATAGAGGACTGTGGAAATTTAAATAATGTTCTTTTCAATTAATCCATCAAGACCA 600

QY 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db 601 TCTATCAAGATTCTGAAAGTCTGTAATCTCAGAAATTAAGTCTGGAGAGACAGTCGCCTTG 660

QY 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db 661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCGAGAGGTTATATGAT 720

QY 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db 721 CCGGATGATGGCTTTATCATGGTGGATGAGAAATGACATCAGAGCTTTAAATGTGGCGCAT 780

QY 398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db 781 TATCGACCATATGGAGTGGTGTAGTCAAGAGCGCTGTTTGTTCGGGACCAACCATCAGT 840

QY 418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluLeuMetGluArgAlaAlaArg 437
Db 841 AACCAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGCAAGG 900
```

```
QY 438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db 901 GAAGCAAAATCGGTATGATTTTATCATGGAGTTTCTTAATAAAATTTAATACATTGGTAGGG 960

QY 458 GluLysGlyValAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db 961 GAAAAAGGAGCTCAATAGTGGAGGCGAGAAACAGAGGATCGCAATTGCTGTGCTTAA 1020

QY 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db 1021 GTTCGAAACCCCAAGATTCTGATTTAGATGAGGTACGCTGCTGCCCTGGATTTCAGAAAGC 1080

QY 498 LysSerAlaValGlnAlaAlaLeuGluLys 507
Db 1081 AAGTCAGCTGTTCAAGCTGCACCTGGAGAAG 1110

RESULT 7
LOCUS AX322791 4189 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 5 from Patent WO0192877.
ACCESSION AX322791
VERSION AX322791.1 GI:18093768
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 Sorrentino,B. and Schuetz,J.
AUTHORS Method of identifying and/or isolating stem cells
TITLE Patent: WO 0192877-A 5 06-DEC-2001;
JOURNAL ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
Source 1. 4189
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 1204 a 875 c 1028 g 1082 t
ORIGIN

Alignment Scores:
Pred. No.: 4.99e-114 Length: 4189
Score: 1590.00 Matches: 307
Percent Similarity: 76.32% Conservative: 99
Best Local Similarity: 57.71% Mismatches: 94
Query Match: 61.34% Indels: 32
DB: 6 Gaps: 5

US-09-873-409-8 (1-514) x AX322791 (1-4189)

QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 148 ATGATTTCTGGAACTCTCGCTGCTATATATCCATGGAACTTACTTCCCTCTTGATGCTG 207

QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAAGCAGAGCCAGTAGTATCTGTCCAAGCATT 267

QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39
Db 268 ACTAATCAAAAGTCGACCCCAACAGTACTCTGATCATCAGCAACAGCAGCTGTGGAGGAAGAG 327

QY 40 -----TyrSerPheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db 328 ATGGCATATACGCCCTAC-----TATTACACCGGGATTGGTGTGGTGTG 372

QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuThrIleThrAlaAlaArgGlnThrLys 76
Db 373 CTCATAGTTGCCCTACATCCAGGTTTCACTTTGCTGCTGTCAGCTGGAGACAGATACAC 432

QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer 96
Db 433 AAGATTAGGCAGGAAGTTTTCATGCTATAATGAATCAGGAGATAGGCTGTTTGTGATGTG 492
```







DTESEKVVQEALDKAREGRTCVIAHRLSTIQNADLIVIEGKVKEGHTHQQLLAQK

GIYFSMVQAGAKRS"

BASE COUNT	1226 a	903 c	1059 q	1110 t

ORIGIN	1220	1	200	3	1000	3
BASE COUNT						

**Alignment Scores:**

Pred. No.:	5.15e-114	Length:	4298
Score:	1590.00	Matches:	307
Percent Similarity:	76.32%	Conservative:	99
Best Local Similarity:	57.71%	Mismatches:	94
Query Match:	61.34%	Indels:	32
DB:	10	Gaps:	5

## US-09-873-409-8 (1-514) x MUSMDR (1-4298)

Qy	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	257	ATGATTCGGGAACCTCGCGTGTATTTATCCATGGAAACATTACTTCCCTCTTGATGCTG	316
Qy	21	ValLeuGlyGluMetSerAspAsnLeuLeu-	30
Db	317	GTGTTTGGAAACATGACAGATAGTTTTCAAAAAGCAGACCCAGTATTCGCCAACGATT	376
Qy	31	-----SerGlyCysLeuValGlnThrAsnThr-----	39
Db	377	ACTAATCAAAAGTGGACCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG	436
Qy	40	-----TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaLa	56
Db	437	ATGGCCATATACGCGCTAC-----TATTACACCGGGATTTGGTCTGCTGTG	481
Qy	57	LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys	76
Db	482	CTCATAGTTGCCATATCCAGGTTTCACTTTGGTGCTCGACGCTGGAGACAGATACAC	541
Qy	77	ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer	96
Db	542	AAGATTAGGCAGGAAGTTTTCATGCTATAATGAATCAGGAGATAGCTGGTTTGATGTG	601
Qy	97	CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly	115
Db	602	CATGATGTGGGGAGCTCAACACCCGGCTCACAGATGATGTCCTCAAAAATTAATGACGA	661
Qy	116	IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla	135
Db	662	ATTGGTGACAAAATTGGGATGTTTTTCAGTGCCATAACCAACATTTTAGCCGGTTTTATC	721
Qy	136	ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle	155
Db	722	ATAGGATTTATAAGTGTGTGGAGCTAAACCTTGTCATTTTGGCTGTCAGCCCTCTTATT	781
Qy	156	MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer	175
Db	782	GGATTGTTCATCTGCTTTGGCGAAAGGTATTGACTTTTACTAATAAGGAACCTCCAG	841
Qy	176	AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle	195
Db	842	GCTTATGCAAAAGCTGGAGAGGTGCTGAAGAAGCTTTAGACCCATCAGACTGTGATT	901
Qy	196	AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys	215
Db	902	GCCTTTGGAGACAACAGAGGAACTTGAAAGGTACATAAATAATTTAGAAGAAGCTAA	961
Qy	216	AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe	235
Db	962	AATGTGGCATAAAGAAGCTATCACAGCCAGCATTTTCGATAGGCATTCGCTACCTGTTG	102
Qy	236	MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu	255
Db	1022	GTCTATGCATCATATGCACCTGGCATTTCTGGTAUGGGAACATCCTTGTGCTCTCAATGAA	108
Qy	256	ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr	275

Db	1082	- - - - - TATTCTATTCGAGAGTGCTTACTGCTCTCTCTCTATTTGTGGGACATTTTT	113
Qy	276	CysIleGlyAlaValProHisPheGluThrPheAlaIleAlaAArgGlyAlaAlaPhe	295
Db	1136	AGTATTGGACACTTGGCCCAACATAGAAGCCTTTGCAAAACGCAGAGGGCGACGCTTT	1195
Qy	296	HisIlePheGlnValIleAspLysLysProSerIleLeuAsnPheSerThrAlaGlyTyr	315
Db	1196	GAAATCTTCAAGATAATTGATACGAGGCCAACGATVTGACAGCTTCTCAACAAGAGGGCTAC	1255
Qy	316	LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer	335
Db	1256	AACCAGACAGTAGTATATGGAAACTTAGAGTATTAANAATGTTCATCTCACTACCCATCG	1315
Qy	336	ArgProSerIleLysIleLeuLysGlyLeuAsnLeuAArgIleLysSerGlyLeuThrVal	355
Db	1316	AGAAGCGAAGTTTCAGATCTTGAAGGSCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG	1375
Qy	356	AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnAArgLeu	375
Db	1376	GCCTTGGTTGGCAACACAGTGCGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC	1435
Qy	376	TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleAArgAlaLeuAsnVal	395
Db	1436	TACGACCCCCTGAGGGCGGTGTCAGTATTCGACGACGAAGACATCAGAACCATCAATGTG	1495
Qy	396	ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr	415
Db	1496	AGGTATCTGAGGGAGATCATTTGGTGTGTGTGATGCAGGAACCTGTGCTGCTTTGCCACCAAG	1555
Qy	416	IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluAArgAla	435
Db	1556	ATCGCCGAGAACATTCGCTATGSCCGAGAGAAGATGTCAACCATGGATGAGATTGAGAAAAGCT	1615
Qy	436	AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu	455
Db	1616	GTCGAAGGAAGCCAATGCCATGATGACTTCATCATGAAACTGCCCCCACTTTTGACACCCCTG	1675
Qy	456	ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnAArgIleAlaIleAlaAArg	475
Db	1676	GTGTTGTGAGAGAGGGCGCGNCTGTAGTGGGGACAGAAACAGAGAAATGCCATTGCCCGG	1735
Qy	476	AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer	495
Db	1736	GCCCTGTGTCGCAATCCCAAGATCCTTTTGTGGAGCAGAGGCCACCTCAGCCCTGTGATACA	1795
Qy	496	GluSerLysSerAlaValGlnAlaAlaLeuGluLys	507
Db	1796	GAAAGTCAAGCTGTGGTCAGGCCGCACTGGATAAG	1831

## RESULT 9

```

AX105082
LOCUS          4279 bp      DNA      linear      PAT 30-APR-2001
DEFINITION     Sequence 26 from Patent WO0123540.
ACCESSION      AX105082
VERSION        AX105082.1  GI:13921232
KEYWORDS
SOURCE         dog.
ORGANISM       Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE      1 (bases 1 to 4279)
AUTHORS       Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
              Patten,C.J.
TITLE          P-glycoproteins and uses thereof
JOURNAL        Patent: WO 0123540-A 26 05-APR-2001;
GENTEST        CORPORATION (US)
FEATURES       Location/Qualifiers
               1..4279
               /organism="Canis familiaris"
               /db_xref="taxon:9615"
               17..3862
               /note="unnamed protein product"
CDS

```

```

/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:1392123"
/translations="MDPEGRKSGAEKNFWMKGSKKKKKKKPKTVSTFAMFRYSN
WLDRLVYLGTMAAI LHGALPLMLVFNGMNTGSFANAGI SRNKTPVLINESI TNNR
QHFNLHBEEMTYAYYSIGAGULVAAYIQVSFWCLAAGRQLIKRKQFPFAIMRW
EIGAFVDVHDVGELNTRLTDDVKSNKEGINGDGKI GMFFQSIAITFTFGIVTRGWKLTL
VILIASPVLGSAAI WAKLISDPTDELLAYAKAGAVAEVLAAIRVTIAFGGQKEL
ERYNKLLEAAGKIGIKAI IANTISIGAFAFLLIYASVALAWFYGTSLVSEYTGIVGL
TFVPSLVLEAGFSIGQASPSTEAFANGAAGAAIEFKIIDNKPISIDTSYKSQSHKPDNIKG
NLFENKVHSPSRKSVKILLKGNLNVQSGQTVALVNSCGCKSTTVOLMRLYDPDT
GMVCIDQDI DRI NVRHREITGVVSGEPVLPFAITTAENRYRGRENVTWDEIEKAVE
ANAYDFIMKLPNFKDTLVGBRGAKLGQQORIAIARALVRNPKILLLLDEATSKLGE
SEAVVQWDPAKARKGRTTIVIAHRLSSTRVNADVIAGDSGGVI VEGKNGHNHEKQGY
FEUVTWOTRGNEILELENATESKSDALEMSDGSSIIKRRTSRHSRIAHPGQDR
KLGTCDENLNVPVPSFWRTL KLNSTWPYFVVIGI PCAI INGGLOPAFSLIFRSRI GII
FTDEPEDENKPNNSVFLPVLGIIISFITFELOGFTFGKAGEILT KLRLRYMVPRSM
LRQVSWDFDPDKMTGALTTRLANDAAQVKAIGSRLAVITQIANLGTGIIISLIGY
WQUTLLLLLPIIAATGVRNWKMLGQDLKKLEGAKGATEIAENRTVVSUSTR
EOXPEYMYAQSLQVYRNLSRKAHIFGVSFSTIQAMMYFSYAGCFRFGALVANEFM
FDQVLVYAIVFECAMAVQGSSVPAKYAKVSAAHVIMI IEKSPILDISYSPHLGRP
NTLEGNVTFNEVYVPTRPDI PVLOGLSLEVKKGTOTALVSGSGCKSTVVOELLERF
YDPLAGSVLIDGKEIKHLNVQWLRAHLGI VQSPILPDCSIAENIAYGDNRSRVVSHEE
IVQAAREANLHHFIETLUPEKYNTRVGDKTQLSGGQQR IAIARALVRQPHELLLEDA
TSQADENSEKVVQVEALDKREGRTICVIAHRLSTI QNADLIVVFQNGKVBKHETHQOOL
LAQKGIYFSMVSVQAQAKR"
BASE COUNT      1296 a   833 c   1009 g   1141 t
ORIGIN

Alignment Scores:
Pred. No.:          1,25e-113           Length:         4279
Score:             1585.00              Matches:        304
Percent Similarity: 75.80%               Conservative:    97
Best Local Similarity: 57.47%            Mismatches:     104
Query Match:       61.15%                Indels:         24
DB:                  6                   Gaps:           4

US-09-873-409-8 (1-514) x AX105082 (1-4279)

Qy  1 MettleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db  167 ATGTTGGTGGGACAATGGCTGCCCATTCATCCATGCAGTGCACCTCCCTCTCATGATGCTG 226

Qy  21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db  227 GTTTTGGAACATGACAGATAGATTTCGAATGCAGGAATTCACGAACCAATTCATCAACCATCTGGAG 286

Qy  35 -----ValGlnThrAsnThyrTyrrSerPheArg----- 44
Db  287 CCAGTTATAATTAATGAAGAATTATACGAACAATATACCAACATTTTCATCAACCATCTGGAG 346

Qy  45 -----LeuThrLeuTyrrTyrrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db  347 GAGGAAAATGACCACGATATGCCTATTATTACAGTGGGATCGGTGCTGCTGGTGGCT 406

Qy  60 GlyTyrlleGinIleSerLeuTriPileIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db  407 GCTTACATCCAGTTTTATTCTGGTGCTGCAGCAGGAGAACAGACATACTCAAATTAGA 466

Qy  80 LysGlnPhePheHisSerValLeuAlaGlnAspileGlyTyrrPheAspSerCysAspile 99
Db  467 AAACAAATTTTTCATGCTATCATCGACAGGAGATTGSGTGGTTTGACGTCATGACGTT 526

Qy  100 GlyGlnLeuAsnThrArgMetThr----AspileAspIlysIleSerAspGlyIleGlyAsp 118
Db  527 GGGGAGCTTAACACCCGGCTCACAGCATGCTTCCAAAAATCAATGAAGGAATTTGGCGAC 586

Qy  119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValcylLeu 138
Db  587 AAAATTGGAATGTCTTTTCAATCAATGACAAATTTTTCACCGGTTTTATATAGTGGGTTT 646

Qy  139 VallysGlyTyrrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158

```

RESULT 10	AX105078	4279 bp	DNA	linear	PAT 30-APR-2001
LOCUS	AX105078	22 from Patent	W00123540.		
DEFINITION	Sequence				
ACCESSION	AX105078				
VERSION	AX105078.1	GI:13921228			
KEYWORDS					
SOURCE	dog.				
ORGANISM	Canis familiaris				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
REFERENCE	1 (bases 1 to 4279)				
AUTHORS	Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.				
TITLE	P-glycoproteins and uses thereof				
JOURNAL	Patent: WO 0123540-A 22 05-APR-2001;				
	GENTEST CORPORATION (US)				
FEATURES	Location/Qualifiers				
source	1..4279				
	/organism="Canis familiaris"				
	/db_xref="taxon:9615"				
CDS	17..3862				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC317721.1"				
	/db_xref="GI:13921229"				
	/translation="MDPEGKGSAEKNFWMGKKSKNKKKKPTVSTFAMPRVSN WLDRLVGLVGTMAAIIHGAAIPLMVLVFGNMTDSFANAGISRNKTPFVINESITNT QHFINLHEEMTTAYVYSGIAGVLVAAYIQVSWCLAAGRQILKIRKQFPAHMRQ EIGWFDVGLGELNTRTDDVSGKIGKIMGFOSIATFTFTLVGFTRGWKLT VILAI5PVLGSAIWAIIKSLSTDPELLAYAKAGAAVEVLAAITRTVAFGGKKEL ERYNKULEAKGIGIKAIITANISIGAAFLIIYASALATWYGTSLVLSSEYISIGVL TFFSVLIGAFSGIQASPSIEAFANARGAAEYIFKIIDNPKSIDSKSGHKPDNKG NLSPKHFYSPRKEVYKIIKGLNKVSGQTVLQNSGCGKSTTVQLMORLYDGD GWICIDQDRIITNVRLHRTIGVSGQVLFATTAENIRYGRENVMTDEIEKAVTE ANAVDFTMKLPNKFDTLVGRGQISGQKORIAARALVRNPKILLDEAT5ALDTE SEAVOVATLDPKARKGRTTIVIAHRLSTVRNADVLGPDGVIIVKGNHDELMEKGYI PKLVMTGRNETELENAATESKSEDALESPKDSGSSLLIKRSTERSIHAQGGDR KLGTKEDLNENVPFWRIKLNSTEWPYFVVGICAIINGGLQPAFSIIFSRIGI FTRDEDPTQRKSNMPSFTVLVLGIIISFTFFLQGTFGKAGEIILTKRLRYMVRSM LRQDVSFDDPTQNTGALTTRLANDAAQVKGAGISRLAVITONIANLGTGIIISLYI WOLTLALLAIVPIIAIAGVNSLRKKAHICGVFSISIQAMMYFSYAGCFRGLVANEFM EOKFVYVAAQSLQVPVYRNSLRKKAHICGVFSISIQAMMYFSYAGCFRGLVANEFM PQDVLVPSAIVGMAAGVQSSFPADPYAKAKYSAAHVIMIEKSPILDSYSPHGLK PNTLEGVTFNVEFVNPTRDIPVLVGLSIVKQGTALVVGSGGCKSTVQVLLERF YDPLAGSLVDIGKEIKHLNVQMLRAHIGIVSQEPIFLDSCIAENIAIARALVRPHILLDEA IMQAKEANTHHIETLPEKYNTRVDGKTQLSGQKORIAARALVRPHILLDEA TSLADTSEKVVQBALDKAREGRTCTVIAHRLSTIQNADLIVVFQNGKVKHEGTHQOOL LAQKGIYFSMVSVOAGAKR"				
BASE COUNT	1295 a	833 c	1008 g	1143 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	2,56e-113	Length:	4279		
Score:	1581.00	Matches:	303		
Percent Similarity:	75.80%	Conservative:	98		
Best Local Similarity:	57.28%	Mismatches:	104		
Query Match:	61.00%	Indels:	24		
DB:	6	Gaps:	4		
US-09-873-409-8 (1-514) x AX105078 (1-4279)					
Qy	1	MettIleLeuGlyIleLeuAlaSerLeuValAnGlyAlaCysLeuProLeuMetProLeu	20		
Db	167	ATGTTGGTGGGCAATGGCTGCCATCATCATGAGTGCACTCCCTCTCATGATGCTG	226		
Qy	21	ValLeuGlyGluMetSerAspAnLeuIleSerGlyCysLeu	34		
Db	227	GTTTGTGGAAACATGACAGATAGCTTTTGCMAATGCAGGAAATTTCAAGAAACAAACTTTT	286		
Qy	35	-----ValGlnThrAsnThyTrSerPhePheArg-----	44		



```

QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaAraGlyAlaAlaPheHisIlePhe 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAAAAGCGCAAGAGGACGCTTATGAAATCTTC 1120

QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1121 AGATAATTGCAATAAACAGCATTCAGCTATTCGAAGATGGACATTAACAGAT 1180

QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1181 AATAATTAGGAAATTGGAAATTCAAATGTTCACTTCAGTTACCTCTCGAAGAA 1240

QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1241 GTTAAGATCTTAAGAGGCTCAACTGAAGGTTACAGTGGCGGACAGTGGCGCTGTT 1300

QY 359 GlyLeuAsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspPro 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1301 GGAACAGTGGCTGCGGGAGAGACGACCGTGCACTGATGAGAGGCTCTATGACCCC 1360

QY 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1361 ACAGATGGCATGCTGTATTGATGGACAGGACATTAGGACCATTAATGTAAGGCATCTT 1420

QY 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1421 CGGAAATTAAGGCTGCTGCTGAGTCAAGGACCTGTGTTGTCACACGACGATGCTGAA 1480

QY 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAraGlu 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1481 ACATTCGCTATGCGCGCGCAATGTCACATGATGATGAGATGAGAAAGCTGTGAAGAA 1540

QY 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1541 GCCAATGCTATGATTTTATCATGAACCTACCTTAATAATTGACACTCTGTTGGAGAG 1600

QY 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1601 AGAGGGCCGACGTAGTGGTGACAGAAACAGAGAATCGCCATTCGCGGCCCTGGTT 1660

QY 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1661 CGCAACCCCAAGATTCCTTCGTGGATGAGGCAACGTCAGCTCGACACTGGAAGTAA 1720

QY 499 SerAlaValGlnAlaAlaLeuGluLys 507
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1721 GCAGTGGTTGAGTGGCCCTGGATAAG 1747

RESULT 12
AX108656 4195 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 3 from Patent WO0123565.
ACCESSION AX108656
VERSION AX108656.1 GI:13923888
KEYWORDS crab-eating macaque.
SOURCE Macaca fascicularis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopithecinae; Macaca.
1 (bases 1 to 4195)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
TITLE P-glycoproteins from macaca fascicularis and uses thereof
JOURNAL Patent: WO 0123565-A 3 05-APR-2001;
GENTEST CORPORATION (US)
FEATURES
source
1. .4195
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
100. .3951
/note="unnamed protein product"
/codon_start=1
CDS
Db 730 TTACACGTGGTGGAAAGCTAACCCCTTGATTTGGCCATCAGTCTCTGTTCTGCACTG 789

```

```

/protein_id="CAC37765.1"
/db_xref="GI:13923889"
/translation="MDLEGDRNGGAEGAKNFKFLANKSKKDKKRRKPTVSVFSMPFRYSN
WLDKLYMVGVTAAAIHHGAGLPLMWLVFGDMTDFPACNAGLGDGLGALLFNNTSSNIT
DTPVKNLEDMTRVAYYSGIGAGLVAAVQVSWFAGLAAAGROIHKTRKGFHAIIMR
QEIWFVDVHGLNRLTDVSKINEGDKIGMFOSMATFFTFVTRFGWKLT
LVIATSPVLGSLAAWAKILSSFTDKELLAYAKAGAAVEVLAIATRVIAFGGOKKE
LERTNKEAKRIGIKAITANISIGAFLLIYASALAFWYGTTLVLSKEYSGQV
LTVFVSVLIGAFSQQQSPSIEAFANARGAAFEIPKIDNKPISIDSYSGHKPNIK
GNLFRNVHFSYPSRKEVKILKLNKLVQSGQTVLVNCGCGKSTTVQLMORLYDPT
EGMVSDGDIRTINVRPLREIIGVVSQEPVLFTATTAENIRYGRDVTMDIEKAVK
EANAVDRTMKLPQKFDTLVGRGAOLSGQKORIAARALVRNPKLILLDEATSLDT
ESEAVQVALDKARKRTIIVIAHRLSVRNADVIAGFDDGVIKEGHDELMEKEGI
IFKLVMTAGNEIELENADESEIDTLEMSSHDSGLIRKRSRHSRVSQGGQD
RKLSTKEALDESIIPVSWFRIMKLNLTWEPVFWVFCAINGGHQAPAFVFSKIIG
IFTRNDATETKRONSLFLLVLGIVSFITFLQGTFGKAGELIKRLRYMVPFRS
MLRODSWFDPPKNTGTALTTLANDAAOVKAIGSLRAITQNTANLGTGIIISLIY
QMOLTLLLAIVPIIATAGVVMKLSQALKDKKELGAGKIACTEENPRFTVVSILT
QEQFEHMYDQSLQVPYRNSLRKAHIFGITFTQMMYFYSGAGCFEGALVAHSLM
SFEDLLVFSAVFGMAVGVSSFPADYAKAKVSAHIIMIEKTPLIDSYSTGLK
PNTLEGNVTNEVFNFTPLDI PVLOGLSLEVKKGOTLALVSGSGCGKSTVQULLR
FYDLAGKVLDDKEIKOLNQLWRAHLGIVSQEILFDCSIENIATGDSNRVVSQE
BIVRAAKEANTHAFIESLPNKYSTRVGDKGTOLSGQKORIAARALVRQPHILLDE
ATSLDTESEKVOEALDKAREGRTCIIVIAHRLSTIQNADLIIVFQNGRVEHGHTHQ
LLAQKGIYFSMVSVQAGAKRQ"
BASE COUNT 1230 a 801 c 1039 g 1125 t
ORIGIN
Alignment Scores:
Pred. No.: 2,74e-113 Length: 4195
Score: 1580.50 Matches: 306
Percent Similarity: 75.85% Conservative: 96
Best Local Similarity: 57.74% Mismatches: 103
Query Match: 60.98% Indels: 25
DB: 6 Gaps: 4
US-09-873-409-8 (1-514) x AX108656 (1-4195)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 ATGGTGGTGGGAACCTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 GTGTTTGGAGACATACGCGATACCTTTTGCAATGTCAGGAAATTTAGGAGATTTAGGAGCT 369

QY 34 LeuValGlnThrAsnThrTyrSer-----41
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 CTGTTTGTAAACACACTAATAGCAGTAAATACATGATACAGTGCCTGATGAATCTG 429

QY 42 -----PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 GAGGAAGATATGACCAAGGTATGCCTATTATTACAGTGGAAATGGTCTGGGTGCTGTT 489

QY 59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 GCTGCTTACATTACGTTTTCATTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAAT 549

QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 AGAAACACGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTATGATGACAGAT 609

QY 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 610 GTTGGGAGCTTAAACCCCGCTTACAGATGATGCTCCAGATTAATAGAGGAATGGT 669

QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 670 GACAAAATTTGGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGA 729

QY 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 730 TTACACGTGGTGGAAAGCTAACCCCTTGATTTGGCCATCAGTCTCTGTTCTGCACTG 789

```





```
Db 329 GTTTTGGAAACATGACAGATAGCTTTGCAAAATGCGAGGAATTTCAAGAAACAAAACTTTT 388
Qy 35 -----ValGlnThrAsnThrTyrSerPhePheArg -----44
Db 389 CCAGTTATAATTAATGAAGTATTAGCAACAATATACAAATTTTCATCAACCATCTGGAG 448
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 449 GAGGAATGACCAACGATATGCTATTATTACAGTGGGATCGGTCTGCGTCTGGTGGCT 508
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 509 GCTTACATCCAGGTTTCATCTGCTGCTCCGCGCAGGAGACAGATATCTCAAAATTAGA 568
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 569 AAACAATTTTTCATGCTATCATGCGACAGAGGATTTGGCTGGTTGACGTGCATGACGTT 628
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 629 GGGGAGCTTAAACACCGCTCACAGACGATGCTCCAAAATCAATGAAGAAATTCGCGAC 688
Qy 119 LysIleAlaLeuLeuPheGlnIleAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 689 AAAATTGGAATGTTCTTTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 748
Qy 139 ValLysGlyTyrPheLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 749 ACACGTGGTGGGAAGTAAACCTTGATTTTGGCATGACGCTGTTCTTGGACATTTCA 808
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 809 GCGGCCATCTGGCAAGATACCTCTCATTTACTGATPAAAGAACTCTTGGCTATGCA 868
Qy 179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTGGA 928
Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 929 GGAAGAAGAAAGAACTTGAAGAGTCAACAAAATTTAGAAAGCTAAAGAAATTTGGG 988
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238
Db 989 ATPAAGAAAGCTATCAGGCCAACATTTCTATTGTTGCTGCTTCTTATTGATCATGCA 1048
Qy 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 1049 TCATATGCTCTGGCTTCTGTTATGGGACCTCTGCTGCTCTCCAGTGAA-----TAT 1102
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1103 TCTATTGGACAAGTACTCACTGCTCTCTTCTGTTGTTATTAATGGGGCTTTTATGATGGA 1162
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePhe 298
Db 1163 CAGGCATCCCAAGCATTGGAAGCATTGGCAACGCAAGAGAGCAGCTTATGAAATCTTC 1222
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1223 AAGATAATTGACAATAAACCACGATTTGACAGCTATTTCGAAGAGTGACATAAACAGAT 1282
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1283 AATATTAAAGGGAATTTGGAATTTCAAAATATGTTTCACCTTCAGTTACCTTCGCGCTGGTT 1342
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db 1343 GTTAAGATCTTAAGGGTCTCAACCTGAGGTTGAGAGTGGGAGACAGCTGGCGCTGGTT 1402
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
```

```
Db 1403 GGAACAGTGGCTGCGGGAAGAGCAGCAGCGTGCAGCTGATGCGAGGCTCTATGACCCC 1462
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1463 ACAGATGGCATGCTGTATTGATGCGACAGGACATTAGGACCAATAATGTAAGGCATCTT 1522
Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db 1523 CGGGAATTAATCTGCTGCTGAGTCCAGGAGCCTGTGTTGTTGCTCCACGATAGTGAA 1582
Qy 419 AsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1583 AACATTGCTATGCGCGCAAAATGTCACCATGATGAGATGAGAAAGCTGTAAGGAA 1642
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1643 GCCAATGCCATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTGGAGAG 1702
Qy 459 LysGlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1703 AGAGGGGCCAGCTGAGTGGTGGCAGAAACAGAGAATCGCCATTCCTCGGGCCCTGGTT 1762
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1763 CGCAACCCCAAGATTTCTCTGCTGGATGAGCAACGTCAGCTCTGGACACTGGAAGTGAA 1822
Qy 499 SerAlaValGlnAlaLeuGluLys 507
Db 1823 GCAGTGGTTTCAGTGGCCCTGGATAG 1849
RESULT 14
AX322787
LOCUS AX322787 3860 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192877.
ACCESSION AX322787
VERSION AX322787.1 GI:18093766
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Sorrentino, B. and Schuetz, J.
TITLE Method of identifying and/or isolating stem cells
JOURNAL Patent: WO 0192877-A 1 06-DEC-2001;
ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES
source 1..3860
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1135 a 746 c 957 g 1022 t
ORIGIN
Alignment Scores:
Pred. No.: 3,23e-113 Length: 3860
Score: 1579.00 Matches: 306
Percent Similarity: 77.04% Conservative: 100
Best Local Similarity: 58.06% Mismatches: 99
Query Match: 60.92% Indels: 22
DB: 6 Gaps: 5
US-09-873-409-8 (1-514) x AX322787 (1-3860)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 151 ATGGTGGTGGGAACATTTTGGCTGCCATCATCCATGGGGCTGGAGACTTCTCTCATGATGCTG 210
Qy 21 ValLeuGlyGluMetSerAsp-----AenLeuIleSer 31
Db 211 GTGTTTGGGAATGACAGATATCTTTGCAATGTCAGGAAATTTAGAGATCTGATGTCGA 270
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-----43
```



```
Db 271 AACATCACTAATAAGAGTGATATCAATGATACAGGGTCTTCTCATGATCTGGAGAGAC 330
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 331 ATGACCAAGATATGCTATTATTACAGTGAATTTGGTCTGGGGTCTGGTTGCTGCTTAC 390
Qy 62 IleglnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 391 ATTCAGGTTTCATTTTGGTCTGGCAGCTGGAAAGACAAATACACAAAATTAGAAAACAG 450
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 451 TTTTTCATGCTATAATCGCAGAGATAGGCTGGTTGATGTGCACGATGTTGGGGAG 510
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 511 CTTAACACCCGACTTTACAGATGATGTCTTAAGATTATGAAGGTATTGGTGACAAATT 570
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 571 GGAATGTCCTTTCAGTCAATGGCAACATTTTCACTGGGTTTATAGTAGGATTTACAGT 630
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 631 GGTGGAGGCTAAACCTTGTGATTTGGCCATCAGTCTGTTCTTGTGACATGTCAGCTGCT 690
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysLeuSerAlaTyrSerLysAla 180
Db 691 GTCGGGCAAGATACATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 751 GGACAGTAGCTGAAGAGCTCTGGCAGCAATTAGACTGTGATGCTTTGGAGGACAA 810
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 811 AAGAAAGAACTTGAAGAGTCAACAAATAATTAGAAAGCTAAAGAAATTTGGGATAAG 870
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 871 AAAGCTATTACGCCAATATTCTATAGTGCTGCTTCTCTGCTGATCTATGCAATCTAT 930
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 931 GCTCTGGCCTTCTGGTATGGACACCTTGTCTCTCAGGGGAA-----TATCTATT 984
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 280
Db 985 GCACAAGTACTCACTGATTTCTTCTGATTAATTTGGGGCTTTTAGTGTGGACAGGCA 1044
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1045 TCTCAAGCATTTGAAGCATTTGCAAAATCGAAGAGAGAGCTTATGAATCTTCAAGATA 1104
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlySerIle 320
Db 1105 ATTGATAATAGCCAGATTTGACAGCTATTGCAAGAGTGGCGACAAACAGATAAATT 1164
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1165 AAGGAAATTTTGAATTCAGAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1224
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1225 ATCTTGAAGGGCCCTGAACCTGAAGGTGAGAGTGGCAGAGCGTGGCCCTGGTTGGAAC 1284
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1285 AGTGGCTCTGGAGAGAGACAAACAGTCCAGCTGATGCAGAGGCCTATGACCCACAGAG 1344
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1345 GGGATGGTTCAGTGTGTCGAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAA 1404
```

```
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1405 ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTTCCACCACGATAGCTGAAACATT 1464
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1465 CGTATGGCGGTGAAATGTCAACATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db 1525 GCTATGACTTTTATCATGAACTGCTCTCATAAATTTCACACCTGTTGGAGAGAGAGGG 1584
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1585 GCCCAGTTGAGTGGTGGCAGAGAGAGGATCGCATTCGACGTGCGCTGTTTCGCAAC 1644
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaIleAspSerGluSerLysSerAla 500
Db 1645 CCCAAGATCCTCTGCTGGATGAGCGCAGCTCAGCTTGGACACAGAAAGCGAAGCATG 1704
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1705 GTTCAGGTGCTCTGGATAAG 1725
```

## RESULT 15

```
E02326          E02326          4378 bp      RNA      linear      PAT 29-SEP-1997
LOCUS          Multidrug resistance relating gene derived from human normal cells.
DEFINITION
ACCESSION      E02326
VERSION        E02326.1  GI:2170561
KEYWORDS       JP 1990100680-A/1.
SOURCE         Homo sapiens.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 4378)
AUTHORS        Ueda,K. and Komano,T.
TITLE          HUMAN NORMAL CELL-DERIVED MDR RELATED GENE
JOURNAL        Patent: JP 1990100680-A 1 12-APR-1990;
SUNTORY LTD
COMMENT         OS Homo sapiens
                PN JP 1990100680-A/1
                PD 12-APR-1990
                PF 05-OCT-1988  JP 1988251475
                PI UEDA KAZUMITSU, KOMANO TORU
                PC C12N15/12,C12N1/21,C12Q1/68;
                CC strandedness: Single;
                CC topology: Linear;
                CC *source: tissue type=Adrenal gland;
                CC *source: clone=SAM1132;
                FH Key Location/Qualifiers
                FH 5'UTR 1..137
                FT CDS 138..3980
                FT /gene='Multidrug resistance relating gene' FT
                mat_peptide 138..3977
                FT /gene='MDRAL'
                FT 3'UTR 3981..>4378.
```

## FEATURES

```
source
Location/Qualifiers
1..4378
/db_xref="taxon:9606"
BASE COUNT 1315 a 818 c 1057 g 1188 t
ORIGIN
Alignment Scores:
Pred. No.: 3.77e-113 Length: 4378
Score: 1579.00 Matches: 306
Percent Similarity: 77.04% Conservative: 100
Best Local Similarity: 58.06% Mismatches: 99
Query Match: 60.92% Indels: 22
```

```
DB: 6 Gaps: 5
US-09-873-409-8 (1-514) x E02326 (1-4378)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAenGlyAlaCysLeuProLeuMetProLeu 20
Db 289 ATGGTGGTGGGAACCTTGGCTGCCATCATCGGTGGGCTGGACTTCCTCTCATGATGCTG 347
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 348 GTGTTGGAGAAATGACAGATATCTTGCAAATGCAGGAAATTTAGAAAGATCTGATGTCA 407
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 408 AACATCACTAATAGAAATGATATCAATGATACAGGCTTCTTCATGAATCTGGAGGAAGAC 467
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 468 ATGACCAGGTATGCCCTATTATTACAGTGAATTTGGTGGGTGCTGGTTCGCTTAC 527
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 528 ATTCAGTTTCATTTGGTGGCTGGCAGCTGGAGACAATACACAATAATTAGAAAACAG 587
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db 588 TTTTTCATGCTATATACGACAGGAGATAGGCTGTTGATGTGCACGATGTTGGGGAG 647
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 648 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAAATGAAGGAATTTGGTACAAAAT 707
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 708 GGAATGTTCTTTCAGTCATAGCAACATTTTCACCTGGGTTTATAGTAGATTACACGT 767
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 768 GGTGGGAAGCTAACCCCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTTCAGCTGCT 827
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 828 GTCTGGGCAAGATACATCTTTCATTTACTGTATAAAGAACTCTTAGCGTATGCAAAAGCT 887
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 888 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTTAGCACTGTGATTCATTTGGAGGACAA 947
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 948 AAGAAAGAACTTGAAGGTACACAAAATTTAGAAGAGCTTAAAGAAATTTGGGATAAAG 1007
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1008 AAGCTATTACAGCCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCATCTAT 1067
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1068 GCTCTGGCTTCTGGTATGGGACCACTTGGTCTCTCAGGGAA-----TATCTATT 1121
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1122 GGACAGTACTCACTGTATTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGGCA 1181
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1182 TCTCCAAGCATTTGAAGCATTTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1241
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1242 ATTTGATAAATGAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCAACCAACAGATAATATT 1301
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
```

```
Db 1302 AAGGGAATTTGGAATTCAGAAATGTTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1361
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1362 ATCTTGAAGGTCTGAACCTGAAGCTGCAGAGTGGGCAGACGCTGCTGCTTGGTGAAC 1421
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1422 AGTGGCTGTGGGAGAGCACACACAGTCCAGCTGTATGCAAGGCTCTATGACCCACAGAG 1481
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1482 GGGATGGTCAAGTGTGTGAGGACAGGATATTAGGACCATAAATGTAAAGGTTTCTACGGAA 1541
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1542 ATCATTTGGTGTGGTGCAGTCAAGAACTGTATTGTTTGGCCACCACGATAGCTGAAAACATT 1601
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1602 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAAGCTGTCAAGGAAGCCAAT 1661
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1662 GCCTATGACTTTTATCATGAACCTGCCTCATAAATTTGACACCTGTTGGAGAGAGGG 1721
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1722 GCCCAGTTGAGTGGTGGGAGAGGAGGATCGCATTTGCACGTGCTGCTGCTGTAAC 1781
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1782 CCCAAGATCCTCTGCTGATGAGGCCACGTGAGCTTGGACACAGAAAAGCAAGCAGTG 1841
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1842 GTTCAGGTGGCTCTGGATAAG 1862
```

Search completed: March 31, 2003, 02:24:15  
JOB time : 7000.72 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 01:20:01 ; Search time 502.349 Seconds  
(without alignments)  
2304.231 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592

Sequence: 1 MITGLIASLVNGACILPLML.....SESKSAVQAALKEKOTPRYSF 514

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cn2\_1/USPTO\_spool/US9873409/runat\_27032003\_115416\_19216/app\_query.fasta\_1.7544  
-DB=N Geneseq 101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US9873409 @CN 1.1 4247 @runat\_27032003\_115416\_19216 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2460.5	94.9	3699	24	ABK83223	Human transporter
2	1590	61.3	4189	21	AAZ49334	Murine multidrug r
3	1590	61.3	4189	24	ABA94367	Mouse BCRP DNA rel
4	1590	61.3	4313	14	AAQ38950	Mouse multidrug re
5	1585	61.1	4279	22	AAD03506	Dog P-glycoprotein
6	1581	61.0	4279	22	AAD03504	Dog P-glycoprotein
7	1581	61.0	4279	22	AAD03505	Dog P-glycoprotein
8	1580.5	61.0	4195	22	AAF86128	Cynomolgous monke
9	1579	60.9	3860	21	AAZ49332	Human wild-type mu
10	1579	60.9	3860	24	ABA94365	Human BCRP DNA rel
11	1576	60.8	4233	21	AAZ90198	Rat mdr1b2 (multis
12	1576	60.8	4233	22	AAF27498	Rat mdr1b2 multidr
13	1576	60.8	4279	22	AAD03488	Dog P-glycoprotein
14	1576	60.8	4317	22	AAD03489	Dog P-glycoprotein
15	1576	60.8	4669	19	AAV32645	Human P glycoprote
16	1576	60.8	4669	24	ABK52041	cDNA encoding huma
17	1572	60.6	4186	22	AAF86127	Cynomolgous monke
18	1571.5	60.6	4254	24	ABK63517	Rat sequence diffe
19	1570	60.6	3860	21	AAZ49333	Human G185V mutat
20	1570	60.6	3860	24	ABA94366	Human BCRP DNA rel
21	1570	60.6	4349	22	AAH57442	Human intestine ce
22	1570	60.6	4646	15	AAQ72872	Human multidrug re
23	1570	60.6	4646	21	AAZ94738	Human ATP binding
24	1570	60.6	4646	24	AAD38994	Human mdr1 gene.
25	1570	60.6	4646	24	ABL68592	Kidney cancer rela
26	1570	60.6	4646	24	ABL68880	Hybrid vector p8F-
27	1570	60.6	6505	17	AAT13394	Retroviral M4 mdr-
28	1570	60.6	8630	21	AAZ24041	Retroviral vector
29	1570	60.6	8630	21	AAZ24042	Mutated human P-g1
30	1568.5	60.5	4264	19	AAV66533	Mutated human P-g1
31	1568.5	60.5	4264	19	AAV66534	Sequence of human
32	1568	60.5	4669	8	AAV70752	Multidrug Resistan
33	1567	60.5	4378	11	AAQ04522	Rat multidrug resi
34	1566.5	60.4	4425	21	AAZ52048	Rat multidrug resi
35	1564.5	60.4	4369	21	AAZ52047	Rat multidrug resi
36	1564	60.3	4669	14	AAQ52726	Sequence of human
37	1562.5	60.3	3840	24	ABL91687	Human polynucleoti
38	1562.5	60.3	3988	21	AAZ88973	Human MDR-1 DNA.
39	1559.5	60.2	4788	21	AAZ49335	Murine multidrug r
40	1559.5	60.2	4788	24	ABA94368	Mouse BCRP DNA rel
41	1479	57.1	3924	21	AAZ94742	Human ATP binding
42	1479	57.1	3924	21	AAZ88974	Human MDR-3 DNA.
43	1479	57.1	3924	24	AEN95801	Gene #2299 used to
44	1445	55.7	3912	24	ABK63653	Rat sequence diffe
45	1273.5	49.1	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

RESULT 1  
ABK83223  
ID ABK83223 standard; cDNA; 3699 BP.  
XX AC ABK83223;  
XX  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.  
XX  
KW Human; ss; gene; transporter and ion channel; TRICH; transport disorder;  
KW neurological disorder; muscle disorder; immunological disorder; cancer;  
KW scleroderma; systemic lupus erythematosus; allergy; leukaemia;  
KW cell proliferative disorder; cervical cancer; breast cancer;  
KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;  
KW myotonic dystrophy; catatonias; endocrine disorder; diabetes;  
KW Grave's disease; gastrointestinal disorder; Crohn's disease;  
KW renal disorder; Good pasture's syndrome; viral infection; cirrhosis;

KW bacterial infection; fungal infection; parasitic infection;  
KW protozoal infection; helminthic infection; cardiovascular disorder;  
KW atherosclerosis, hepatic disease.  
XX Homo sapiens.  
OS  
XX WO200240541-A2.  
PN  
XX 23-MAY-2002.  
PD  
XX  
PF 25-OCT-2001; 2001WO-US46055.  
PR  
XX 27-OCT-2000; 2000US-243989P.  
PR 03-NOV-2000; 2000US-245904P.  
PR 09-NOV-2000; 2000US-247673P.  
PR 17-NOV-2000; 2000US-249661P.  
PR 20-NOV-2000; 2000US-252232P.  
PR 01-DEC-2000; 2000US-250790P.  
XX  
XX (INCYTE GENOMICS INC.  
PA  
XX Tang YT, Yue H, Nguyen DB, Hafalia AJA, Elliott VS, Lu Y;  
PI Wallia NK, Yao MG, Baughn MR, Gandhi AR, Ding L, Sanjanwala M;  
PI Ramkumar J, Arvizu C, Gietzen KJ, Lal PG, Azimzai Y, Khan FA;  
PI Thangavelu K, Thornton M, Lu DAM, Tribouley CM, Warren BA;  
PI Ison CH, Das D, Raumann BE, Policky JL, Kearney L;  
XX  
XX WPI; 2002-463570/49.  
DR P-PSDB; ABG61544.  
XX  
XX New transporters and ion channels (TRICH) polypeptides, useful for  
PT diagnosing, preventing, and treating disorders associated with an  
PT abnormal expression or activity of TRICH, e.g. immunological, muscular  
PT or renal disorders  
XX  
XX Claim 5; Page 172; 178pp; English.  
XX  
XX The invention relates to human transporters and ion channels (TRICH)  
CC polypeptides, a naturally occurring amino acid sequence 90 % identical to  
CC TRICH, a biologically active fragment of TRICH or an immunogenic fragment  
CC of TRICH. Also included are an isolated polynucleotide encoding TRICH,  
CC a recombinant polynucleotide comprising a promoter sequence operably  
CC linked to the TRICH polynucleotide, a cell transformed with the  
CC recombinant polynucleotide, a transgenic organism comprising the  
CC recombinant polynucleotide, an isolated antibody that binds specifically  
CC to TRICH, and screening for compounds which bind to TRICH, modulate  
CC TRICH, modulate TRICH expression or are ant/agonists of TRICH.  
CC The polypeptides are useful for diagnosing, treating, and  
CC preventing transport, neurological, muscle, immunological disorders  
CC (e.g. scleroderma, systemic lupus erythematosus, allergies), cell  
CC proliferative disorders such as cancers (e.g. leukaemia, cervical or  
CC breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,  
CC Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,  
CC catatonla), endocrine disorders (e.g. diabetes, Grave's disease),  
CC gastrointestinal disorders (e.g. Crohn's disease), renal disorders  
CC (e.g. Good pasture's syndrome), vital, bacterial, fungal, parasitic,  
CC protozoal and helminthic infections, cardiovascular disorders (e.g.  
CC atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many  
CC other diseases and disorders detailed in the specification. They can also  
CC be used in assessing the effects of exogenous compounds on the  
CC expression of nucleic acid and amino acid sequences of transporters and  
CC ion channels. TRICH or its fragments may also be used in screening for  
CC compounds that specifically bind to and modulate the activity of TRICH.  
CC The polynucleotides can be used to create knock-in humanised animals or  
CC transgenic animals to model human disease. The present sequence  
CC encodes a TRICH protein.  
XX  
XX Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:  
Pred. No.: 1.53e-246 Length: 3699  
Score: 2460.50 Matches: 495  
Percent Similarity: 95.40% Conservative: 3

Best Local Similarity: 94.83% Mismatches: 9  
Query Match: 94.93% Indels: 15  
DB: 24 Gaps: 2  
US-09-873-409-8 (1-514) x ABK83223 (1-3699)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 145 ATGATCCTGGGTATACCTGACATCAGTCTGTTCAATGGAGCCTCCCTTCTTAATGCCACTG 204  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
DB 205 TGTATAGGAGAAATAGTGATACCTTATTAGTGGATGCTTAGTCCACACACACAACA 264  
QY 41 SerPhePheArg-----LeuThr 46  
DB 265 AATTATCAGAACTGTACTCAGTCTCAAGAGAGAGCTGAATGAAGATATGACTCTGTTGACC 324  
QY 47 LeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
DB 325 CTGTATTATGTTGGAATAGTGTGCTGCTTGGTTTGGTTTACATACAGATTCTCTTG 384  
QY 67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleAargLysGlnPhePheHisSerVal 86  
DB 385 TGGATTATTAACCTGCAGCAGCAGACAGAGAGATTTCGAAACAGATTTTTCATTAGTT 444  
QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
DB 445 TTGGCAGCAGACATCGCTGGTTTGTAGTGTGATCTCGTGAACCTTAACTCTCGCATG 504  
QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
DB 505 ACAGATGACATTGACAAATCAGTGTATGTTGGAGATAAGATTCTCTGTTGTTTCAA 564  
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145  
DB 565 AACATGCTACTTTTTTCGATTGGCTGGCAGTTGGTGTGGTGAAGGGCTGGAAACCTCACC 624  
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165  
DB 625 CTAGTGACTCTATCAGCTCTCTCTTATTAATGGCTTCAGCGCAGCATGTTCTAGATG 684  
QY 166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185  
DB 685 GTCATCTCATTTGACAGTAAAGAAATTAAGTGCCTATTCCAAAGCTGGGGCTGGCAGAA 744  
QY 186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGln 205  
DB 745 GAAAGTCTTGTCATCAATCCAGACAGTCATAGCCTTTAGGGCCCGAGGAAAGAACTTCAA 804  
QY 206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225  
DB 805 AGGTATACAGAGATCTCAAGATGCAAGAGATTTGGGCATATAAAGAGACTATAGCTTCA 864  
QY 226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245  
DB 865 AAAGTGCTCTTGCTGTGTACTTCTTTATGAATGAACCTTATGAGCTTCTTTTGG 924  
QY 246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265  
DB 925 TATGGAAACCTCTCTTGATTCTTATGGAGAACTCGGATATACCATCGGGAGCTGTCTTGTCT 984  
QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
DB 985 GTTTTCTTTTAGTGTATTCATAGTATTGTATTTGGAGCAGCAGTCCCTCTACTTTGAA 1044  
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305  
DB 1045 ACCTTCGCAATACCCCGAGGAGCTGCTTTCATATTTTCCAGGTATTATGATAAGAACCC 1104  
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325  
DB 1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAGGAACCTGTGGAA 1164

QY 326 PheLYAsnValSerPheAsnTyrProSerArgProSerIleLysLeuLYsGlyLeu 345  
DB 1165 TTTAAAAATGTTCTTCAATTATCCATCAAGCCATCTATCAAGATTCGAAAGGCTCG 1224  
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLYs 365  
DB 1225 AATCTCGGAATTAAGTCTCGAGAGACAGTCGCTTGGTCGCTCAATGCGACGTGGGAAG 1284  
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385  
DB 1285 AGTACGGTAGTCAGCTTCTGCAGAGGTTATATGATCGGATGATGGCTTATCATGGTG 1344  
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405  
DB 1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTTATCGAGACCATATGGAGTGGTT 1404  
QY 406 SerGlnGluProValLeuPheGlyThrIleSerAsnAsnIleLysTyrGlyArgAsp 425  
DB 1405 AGTCAAGAGCCTGTTTGTTCGGGACCATCATCAGTAACAATATCAAGTATGACGAGAT 1464  
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445  
DB 1465 GATGTGACTGATCAAGAGATGGAGAGCGACGACGAGGAACCAATGCGTATGATTTTATC 1524  
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSergly 465  
DB 1525 ATGGAGTTTCTTAATAATTAATACATTGTTAGGAGGAAAAAGAGCTCAATGAGTGGGA 1584  
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485  
DB 1585 GGCGAGAAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTTCGAAACCCCAAGATTCGATT 1644  
QY 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505  
DB 1645 TTAGTAGGCTACGCTGCTCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACGTG 1704  
QY 506 GluLYs 507  
DB 1705 GAGAAG 1710  
RESULT 2  
ID AAZ49334  
XX AAZ49334;  
AC AAZ49334;  
XX  
DT 14-MAR-2000 (first entry)  
XX  
DE Murine multidrug resistance-1 (MDR-1) cDNA.  
XX  
KW Multidrug resistance; MDR-1; P-glycoprotein;  
KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
KW gene therapy; gene replacement; genetic defect; thalasassaemia;  
KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
KW cytokine; ds.  
XX  
OS Mus sp.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3831  
FT /tags a  
FT /product= "Murine MDR-1 protein"  
XX  
XX WO961589-A2.  
XX  
XX  
PD 02-DEC-1999.  
XX  
XX 27-MAY-1999; 99NO-US11825.  
PF  
XX 28-MAY-1998; 98US-0086988.  
PR  
XX

(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
Sorrentino B, Bunting K;  
WPI: 2000-072615/06.  
P-PSDB; AAY58188.  
Ex vivo expansion of hematopoietic stem cells transduced with a  
PT sequence encoding human multidrug resistance-1, used for bone marrow  
PT transplantation -  
XX  
PS Disclosure; Page 90-93; 113pp; English.  
XX  
CC This sequence represents cDNA encoding murine multidrug  
CC resistance protein MDR-1. MDR-1 is a transmembrane  
CC efflux pump, responsible for the export of drugs from cells,  
CC particularly cancer cells. The invention relates to transducing  
CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified haematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC haematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from haematopoietic stem cells, e.g., thalasassaemia,  
CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in haematopoietic stem cells.  
CC Haematopoietic stem cells that express MDR-1 will be protected against  
CC chemotherapeutic agents, so can be engrafted while the patient is  
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
CC provides sufficient cells to permit standard biochemical analysis.  
CC Overexpression of MDR-1 allows cytokine-driven expansion of  
CC haematopoietic stem cells by at least 10-fold compared with a maximum  
CC of 4-fold in known procedures.  
XX  
SQ Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 1.74e-155 Length: 4189  
Score: 1590.00 Matches: 307  
Percent Similarity: 76.32% Conservative: 99  
Best Local Similarity: 57.71% Mismatches: 94  
Query Match: 61.34% Indels: 32  
DB: 21 Gaps: 5  
  
US-09-873-409-8 (1-514) x AAZ49334 (1-4189)  
  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 148 ATGATTCTGGGAACCTCTCGCTGCTATTATCCATGGAACTTACTCCCTCTTGTGCTG 207  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
DB 208 GTGTTTGGAAACATGACATAGTTTACAAAGCAGAGCCAGTATTCTGCCAAGCAT 267  
QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
DB 268 ACTAATCAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327  
QY 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAla 56  
DB 328 ATGGCCATATACGCTAC-----TATTACCCGGGATGGTGTGTGTG 372  
QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76  
DB 373 CTATAGTTGCGCTACATCCAGGTTTCACATTGTGCTGGCAGCTGGAGACAGATACAC 432  
QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
DB 433 AAGATTAGGCAGAGATTTTTCATGCTATAATGAATCAATCAGAGATAGCTGGTTGATGTG 492  
QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGly 115  
DB 493 CATGATGTTGGGAGCTCAACACCCCGCTCACAGATGATGTCTCCAAAAATTAATGACGGA 552





DB: 24 Gaps: 5

US-09-873-409-8 (1-514) x ABA94367 (1-4189)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 148 ATGATTTCTGGAACTCTCGCTGCTATTATTCATGGAAACATTACTTCCCTCTCTGATGCTG 207  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30  
Db 208 GTCTTTGGAAACATGACAGATAGTTTACAAAAGCAGAACCCAGTATTCTGCCAAGCATT 267  
Qy 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
Db 268 ACTAATCAAAAGTGGACCCCAACAGTACTCTGATCATCAGCAACGACGATCTGGAGGAAGAG 327  
Qy 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAla 56  
Db 328 ATGCCATATACGCCCTAC-----TATTACACCGGATTTGGTGTGCTGTG 372  
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76  
Db 373 CTCATAGTTGCCATACATCCAGGTTTCACTTTGGTGCCTGCGCAGCTGGAACAGATACAC 432  
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
Db 433 AAGATTAGGCAGGAAGTTTTCATGCTGCTATTAATGAATCAGGAGATAGGCTGGTTGATGTG 492  
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115  
Db 493 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATCTCCAAAATAATGACGGA 552  
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
Db 553 ATTGGTGACAAATGGGATGTTTTTTCAGTCCATAACCAACATTTTAGCCGGTTTTATC 612  
Qy 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
Db 613 ATAGGATTTTATAAGTGGTGGAGCTAAACCTTGTCTCATTTTGGTGTGTCAGCCCTCTTATT 672  
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
Db 673 GGATTGTCATCTGCTTGTGGCAAGAGTATTGACTTCATTACTATAGGAACCTCCAG 732  
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195  
Db 733 GCTATTGCAAAAGCTGGAGCAGTGTCTGAGAAAGTCTTAGCAGCCATCAGAACTGTGATT 792  
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215  
Db 793 GCCTTTGGAGGACACAGAGGAAGCTTGAAGGTACAAATAAAATTTAGAAGAAGCTAAA 852  
Qy 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235  
Db 853 AATGTTGGCATAAAGAAAGCTATACAGCCAGCATTTTCGATAGGCATTTGCTACCTGTTG 912  
Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGlu 255  
Db 913 GTCTATGCATCATATGACATGCGATTCGTGTATGGGACATCTCTGGTCTCTCAATGAA 972  
Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275  
Db 973 -----TATTCTATTGGAGAAGTGTCTTACTGTCTCTCTCTATTATTGTTGGGACTTTT 1026  
Qy 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295  
Db 1027 AGTATTGGACATTTGGCCCAACATAGAGACCTTTTCCAAACGACGAGGGCGAGCCTTT 1086  
Qy 296 HistIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315  
Db 1087 GAAATCTTCAAGATAATTGATAACGACGACGACGATTCGACGCTTCTCAACAAAGGCGTAC 1146  
Qy 316 LysProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335

Db 1147 AAACACAGACAGTATAATGGAAACCTTAGAGTTTAAAAATGTTTCACTTCACTACCATCG 1206  
Qy 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355  
Db 1207 AGAAGCGAAGTTTCAGATCTTGAAGGCGCTCAATCTGAAGGTGAAGCGGACAGCGGTG 1266  
Qy 356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeu 375  
Db 1267 GCCTTGGTGGCAACATGCTGTGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1326  
Qy 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395  
Db 1327 TAGACCCCTCGAGGCGCTGTGTCAGTATCGACGGCAAGACATCAGAACCATCAATGTG 1386  
Qy 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415  
Db 1387 AGTATCTCAGGAGATCATTTGGTGTGTGAGTCAGAACTGTGCTGTTGCCACACG 1446  
Qy 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluMetGluArgAla 435  
Db 1447 ATCGCGAGAACATTCGCTATGCGCCGAGAAGATGTCCATGATGAGATTGAGAAAGCT 1506  
Qy 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455  
Db 1507 GTCAGGAAGCCAAATGCCATTCATCATGAAACTGCCCACTTTGACACCCCTG 1566  
Qy 456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475  
Db 1567 GTTGTGAGAGAGGGCGGAGCTGAGTGGGGACAGAAACAGAGAAATCGCCATGCGCGG 1626  
Qy 476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495  
Db 1627 GCCTGTGTCGCAATCCCAAGATCCTTTTGTGGACGAGGCCACCTCAGCCCTGGATACA 1686  
Qy 496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507  
Db 1687 GAAAGTGAAGCTGTGTGCGGCCGCTGGATAAG 1722

RESULT 4  
AQ038950  
ID AQ038950 standard; DNA; 4313 BP.  
XX  
AC AQ038950;  
DT  
XX 28-JUL-1993 (first entry)  
DE Mouse multidrug resistance sequence.  
XX  
KW mdr gene; Lambda DR11 clone; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 110..3940  
FT /\*tag= a  
FT /phenotype= multidrug\_resistance  
XX  
PN US5198344-A.  
XX  
PD 30-MAR-1993.  
XX  
XX  
PF 15-JUL-1986; 86US-0885951.  
XX  
PR 15-JUL-1986; 86US-0885951.  
PR 06-FEB-1991; 91US-0652311.  
XX  
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.  
XX  
XX Croop JM, Gros P, Houseman DE;  
XX  
XX WPI; 1993-126077/15.  
DR P-PSDB; AAR35199.

XX DNA sequence which confers multi-drug resistance on sensitive  
PT mammalian cells - used to preserve bone marrow cells during  
PT chemotherapy to prevent infection  
XX  
XX Claim 1; Fig 8; 22pp; English.  
XX  
XX A cDNA library was constructed from mRNA which had been isolated  
CC from a drug-sensitive mouse cell-line. Two mouse cDNA molecules  
CC complementary to the mRNA species encoded by 2 related but distinct  
CC mdr genes were isolated and cloned. One of the cDNA clones (lambda  
CC Dril) is a full-length cDNA clone for one member of the mdr gene  
CC family. When incorporated into prokaryotic expression vector pDREX4  
CC (which allows high levels of transcription of the cDNA when  
CC introduced into mammalian cells), the clone was shown to confer the  
CC multidrug resistance phenotype upon transfection into drug-sensitive  
CC mammalian cells.  
XX  
SQ Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:

Pred. No.:	1.81e-155	Length:	4313
Score:	1590.00	Matches:	307
Percent Similarity:	76.32%	Conservative:	99
Best Local Similarity:	57.71%	Mismatches:	94
Query Match:	61.34%	Indels:	32
DB:	14	Gaps:	5

US-09-873-409-8 (1-514) x AAQ38950 (1-4313)

QY 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 257 ATGATTCTGGGAATCTCGCTGCTATTATCCATCGAATCTTCCCTCTTGATCGTG 316

QY 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30  
DB 317 GTGTTTGGMAACATGACAGATAGTATTTTAAAGAGCAGAGCCAGTATCTGCGCAAGCATT 376

QY 31 -----SerGlyCysLeuValGlnThrAsnThr----- 39  
DB 377 ACTAATCAAGTGGACCCACAGACTCTGATCATCGACCAACAGCAGCTCGGAGGAAGAG 436

QY 40 -----TyrSerPhePheArgLeuThrLeuTyrValGlyIleGlyValAlaAla 56  
DB 437 ATGCCCATATAGGCTAC-----TATTACACCGGATGTGCTGCTGTG 481

QY 57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76  
DB 482 CTCATAGTTGCTACATCCAGGTTTCACTTTGTGCTGCGCAGCTGGAAGACAGATACAC 541

QY 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96  
DB 542 AAGATTAGCGCAAGATTTTCCATGCTATATGAATCAGGAGATAGCTGTGTTGATGTG 601

QY 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115  
DB 602 CATGATGTTGGGAGCTCAACACCGGCTCACAGATGATGCTCCAAAATTAATGACGGA 661

QY 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
DB 662 ATTGCTGACAAAATGGGATGTTTTTTTCAGTCCATAACCAATTTTAGCGCGTTTATC 721

QY 136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
DB 722 ATAGATTTTATAAGTGGTGGAACTAACCTTGCTCATTTTGGCTGTCAGCCCTCTTATT 781

QY 156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
DB 782 GCATTGTCATCTGTTTGGGCAAGGATTTGACTTCACTTACTATTAAGAACTCCAG 841

QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195  
DB 842 GCTTATGCAAAAGCTGGAGCAGTGTGCTGAAGAGTCTTAGCGCCATCAGAACTGTGATT 901

QY 196 AlaPheArgAlaGlnGlnLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215  
DB 902 GCCTTTGGAGGACAAACAGAGGAACTTGAAGGTACAATAAAATTTAGAAAGACTAAA 961

QY 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235  
DB 962 AATGTTGGCATAAAGAAAGCTATACAGCAGCATTTTCGATAGGCATTCCTACCTGTTG 1021

QY 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrGlyThrSerLeuIleLeuAsnGlyGlu 255  
DB 1022 GTCATATGCATCATATCATCTGGCATTCGTGTATGGACATCTCTGTCTCTCAATGAA 1081

QY 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyr 275  
DB 1082 -----TATCTATTCGAGAGAGTCTTACTGTCTCTCTCTATTTTGTGGGGACTTTT 1135

QY 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295  
DB 1136 AGTATTGGACACTTGGCCCAACATAGAGCCTTTTGCAAACGACGAGGGGAGCCTTT 1195

QY 296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315  
DB 1196 GAAATCTTCAAGATAATTGATAACGAGCCAAAGCAATTTGACAGCTTCTCAACAAAGGGCTAC 1255

QY 316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335  
DB 1256 AAACCCAGACAGTATATATGGAACTTAGAGTTTAAATAATGTTCACTTCACTACCATCG 1315

QY 336 ArgProSerIleLysIleLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355  
DB 1316 AGAAGCGAAATTCAGATCTTGAAGGCGCTCAATCTGAAGGTGAAGAGCGGACAGCGTG 1375

QY 356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375  
DB 1376 GCCTTGTGTTGGCAACAGTGGCTGTGGAAGAACACAACTGTCTCCAGCTGATGCGAGAGCTC 1435

QY 376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395  
DB 1436 TAGACCCCTCGAGGGCGGTGTCAGTATCGACGGACAGACATCAGAACCATCAATGTG 1495

QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415  
DB 1496 AGTATCTGAGGCGAGATCATTTGTGTGTGAGTCAGGAACCTGTGCTGTTGCCACACG 1555

QY 416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435  
DB 1556 ATCGCCGAGAACATTCGCTATGCGCCGAGAGATGTCAACCATGATGATGATGAGAAGCT 1615

QY 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455  
DB 1616 GTCAGGAAGCCAAATGCTATGATCTTCACTGAAACTGCCCAACCAATTTGACCCCTG 1675

QY 456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475  
DB 1676 GTTGTGTGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAAATCGCAITTCGCGG 1735

QY 476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495  
DB 1736 GCCTGTGTCGCAATCCCAAGATCCTTTTGTGGAGGAGGCCCTCAGCCCTGGATACA 1795

QY 496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507  
DB 1796 GAAAGTGAAGCTGTGTGTCAGGCGGCAGCTGGATAG 1831

RESULT 5  
AAD03506  
ID AAD03506 standard; cDNA; 4279 BP.  
XX  
AC AAD03506;  
XX  
DT 13-JUN-2001 (first entry)  
XX

DE Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.  
 XX  
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KW MDRI; drug bioavailability; transgenic animal; genetic model, ss.  
 XX  
 OS Canis familiaris.  
 XX  
 FH Location/Qualifiers  
 FT CDS 17..3862  
 FT /tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype D) protein"  
 FT replace (91, T)  
 FT allele  
 FT /tag= b  
 FT replace (607, C)  
 FT allele  
 FT /tag= c  
 FT replace (1001, T)  
 FT allele  
 FT /tag= c  
 FT replace (3458, A)  
 FT allele  
 FT /tag= c  
 XX WO200123540-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 XX 28-SEP-2000; 2000WO-US26767.  
 PF  
 XX 28-SEP-1999; 99US-0156510.  
 PR  
 XX (GENT-) GENTEST CORP.  
 PA  
 XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
 PI  
 XX WPI: 2001-235373/24.  
 DR  
 XX P-PSDB; AAE00310.  
 DR  
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 PT  
 XX Claim 9; Page 102-107; 11pp; English.  
 PS  
 XX The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDRI) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein, PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype D) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 XX  
 SQ Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5,96e-155 Length: 4279  
 Score: 1585.00 Matches: 304  
 Percent Similarity: 75.80% Conservative: 97  
 Best Local Similarity: 57.47% Mismatches: 104  
 Query Match: 61.15% Indels: 24  
 DB: 22 Gaps: 4  
 US-09-873-409-8 (1-514) x AAD03506 (1-4279)  
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 167 ATGTTGGTGGGACAAATGGCTGCATCATCCATGAGCTGCACTCCCTCTCATGATGCTG 226  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34  
 Db 227 GTTTTGGGAACATGACAGATAGCTTTGCAAAATGCAGGAATTTCAAGAAACAAACTTTT 286  
 QY 35 -----ValGlnThrAsnThrThrSerPhePheArg----- 44  
 Db 287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACTTTCATCAACCATCTGGAG 346  
 QY 45 -----LeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePhe 59  
 Db 347 GAGGAAATGACCACTATGCTTATTACAGTGGGATCGGTGGCGTGTGGTGGCT 406  
 QY 60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79  
 Db 407 GCTTACATCCAGTTTTCATCTGGTGGCTGGCAGCAGGAAGACAGATCTCAAAATTAGA 466  
 QY 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99  
 Db 467 AAACAATTTTTCATGCTATCATGCGACAGAGATTGGCTGTTTACGTGATGACGTT 526  
 QY 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118  
 Db 527 GGGGAGCTTAACACCCGGCTCACAGCATGTCTCAAATCAATGAGGAATTGGCGAC 586  
 QY 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138  
 Db 587 AAATTTGGNATGTTCTTTCATATCAATGACCAACTTTTCCCGGTTTATAGTGGGTTT 646  
 QY 139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158  
 Db 647 ACACGTGGTGGAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACATTCA 706  
 QY 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178  
 Db 707 GCGGCATCTGGGCAAGATATCTTCTTACTTACTGATAAAGAACTCTTGGCCTATGCA 766  
 QY 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198  
 Db 767 AAAGCTGGAGCAGTAGCTGAAAGAGTCTTAGCAGCAATCAGAACTGTGATTCCTTGA 826  
 QY 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218  
 Db 827 GCACAAAAGAAAGAACTTGAAGGTACAAACAAAATTTAGAAAGCTAAAGAAATTGGG 886  
 QY 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238  
 Db 887 ATAAAGAAAGCTATCAGCCCAACATTTCTATTTGGTGGCGCTTCTTATGTATCTATGCA 946  
 QY 239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258  
 Db 947 TCATATGCTCTGGCTTCTGTATGGACCTCTGTGTCTCTCCAGTGAA-----TAT 1000  
 QY 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278  
 Db 1001 ACTATTGGACAAGTACTCACTGCTCTTCTTCTGTATTAAATTTGGGGCTTTTAGTATTGA 1060  
 QY 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298  
 Db 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120  
 QY 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318  
 Db 1121 AAGATAATTGACAATAAACAAGCTATGACAGCTATTGAAAGAGTGGACATAAACAGAT 1180  
 QY 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338  
 Db 1181 AATATTAGGGAATTTGGAATTCAAATTTCACTTCAGTTACCTTCCAGTAAAGAA 1240  
 QY 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358



```
Db 707 GCGCCATCTGGCAAGATACCTATCTTCACTGATTAAGAACTCTGGCGCTATGCA 766
Qy 179 LyeAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGCTTAGCAGCAATCAGAAGCTGTGATGCTTTGGA 826
Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 827 GGCACAAAGAAAGAACTTGAAGGTACACAAATAATTTAGAAGAAGCTAAAGAAATGGG 886
Qy 219 IleLysArgThrIleAlaSerLysValSerLysGluValValTyrPhePheMetAsnGly 238
Db 887 ATAAAGAAAGCTATACGCGCAACATTTCTATTGGTCCGCTTCTATTGATCATGCA 946
Qy 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 947 TCATATGCTCTGGCTTCTGATGGGACCTCTCTGGTCTCTCCAGTGAA-----TAT 1000
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1001 TCTATTGGACAAGTACTCACTGCTCTTCTTCTGTTATTAATTTGGGGCTTTTAGTATTGGA 1060
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db 1061 CAGCATCCCCAGCATTGAAGCATTTGCAACGCAAGAGAGAGCAGCTTATGAATCTTC 1120
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1121 AAGTATTGCAATAAACCAAGCATTTGACAGCTATTGGAAGAGTGGACATAAACAGAT 1180
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1181 AATATTAAAGGGAATTTGGAATTCAAAATGTTCACTTCACTTACCTTCTCGAAAGAA 1240
Qy 339 IleValIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db 1241 GTTAAGATCTTAAGGGTCTCAACCTGAAGGTCAGAGTGGGAGAGTGGCGCTGGTT 1300
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db 1301 GGGACAGTGGCTGGGGAAGAGACACGCGTGCAGCTGATGCAGAGGCTCTATGACCC 1360
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1361 ACAGATGGCATGCTGTATTGATGGACAGACATTAGGACCAATAATGTAAGGCATCTT 1420
Qy 399 ArgAspHisIleGlyValValSerGlnProValLeuPheGlyThrThrIleSerAsn 418
Db 1421 CGGGAATTAATCTGTTGGTGGTGCAGAGCCCTGTGTGTTGCCACCATAGCTGAA 1480
Qy 419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1481 ACATTCGCTATGGCGCGAAATGTCACCATGATGATGAGATTGGAAGCTGTTAAGAA 1540
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1541 GCCAATGCCTATGATTTTATCATGAACACTACCTAATAAATTTGACACTCTGGTTGGAG 1600
Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1601 AGAGGGGCCAGCTGAGTGTGTGACAGAAACAGAGAAATCGCCATTGCTGGGGCCCTGGTT 1660
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1661 CGCAACCCCAAGATTTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACATGAAAGTGAA 1720
Qy 499 SerAlaValGlnAlaAlaLeuGluLys 507
Db 1721 GCAGTGGTTCCAGGTGGCGCTGGATAAG 1747
RESULT 7
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.
```

```
XX AAD03505;
XX 13-JUN-2001 (first entry)
XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
XX MDRI; drug bioavailability; transgenic animal; genetic model; ss.
XX Canis familiaris.
XX Key Location/Qualifiers
XX CDS 17..3862
XX /tag= a
XX /product= "Dog P-glycoprotein (PGP) allelic variant
XX (Genotype B) protein"
XX allele replace (91, T)
XX /tag= b
XX allele replace (607, C)
XX /tag= c
XX WO200123540-A2.
XX PN
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26767.
XX 28-SEP-1999; 99US-0156510.
XX (GENT-) GENTEST CORP.
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI: 2001-235373/24.
XX P-PSDB; AAE00309.
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX Claim 9; Page 93-99; 11pp; English.
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDRI) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.
XX Sequence 4279 BP; 1296 A; 833 C; 1142 T; 0 other;
Alignment Scores:
Pred. No.: 156e-154 Length: 4279
Score: 1581.00 Matches: 303
Percent Similarity: 75.80% Conservative: 98
Best Local Similarity: 57.28% Mismatches: 104
Query Match: 61.00% Indels: 24
DB: 22 Gaps: 4
US-09-873-409-8 (1-514) x AAD03505 (1-4279)
```

```
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 167 ATGTTGGTGGGACAATAGGCTGCCCATCATCCATGAGCTCCCTCTCATGATCGT 226
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----- 34
Db 227 GTTTTGGAAACATGACAGATAGCTTTGCAAAATCAGGAATTTCAAGAAACAAAACCTTTT 286
Qy 35 -----ValGlnThrAsnThrTyrSerPheArg----- 44
Db 287 CCAGTTAATAATGAAGATTATGCAACAATACACAATTTCAACACCATCTGGAG 346
Qy 45 -----LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db 347 GAGGAATGACCACGTATGCTATTATACAGTGGATCGGTGCGCTGCTGGTGGCT 406
Qy 60 GlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db 407 GCTTACATCCAGGTTTCATTCTGGTCTGCGCAGCAGGAAGACAGATACTCAAAATTAGA 466
Qy 80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIle 99
Db 467 AAACAAATTTTTCACTATCATCGCAGGAGATTGCTGTTGACGTGCATGACGTT 526
Qy 100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db 527 GGGGAGCTTAACACCGGCTCACAGACGATGCTCCAAAATCAATGAAGAAATTGGCGAC 586
Qy 119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db 587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTCCACCGGTTTTATAGTGGGGTTT 646
Qy 139 ValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db 647 ACACGTGGTGGAAAGCTAACCTTGTGATTTGGCCATCAGCCCTGTTCTTGGACTTCA 706
Qy 159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db 707 GCCGCCATCTGGGCAAGATACTATCTTCAATTTACTATAAAGAACTCTTGGCCTATGCA 766
Qy 179 LysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db 767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTCCTTTTGG 826
Qy 199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db 827 GGACAAAGAAAGAACTTGAAGGTACAACAAAATTTAGAAGAACTGAAGAAATTGGG 886
Qy 219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPheMetAsnGly 238
Db 887 ATAAAGAAAGCTATCACGGCCCAACATTTCTATTGGTGGCGCTTTCTTATTGATCTATGCA 946
Qy 239 ThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db 947 TCATATGCTCTGGCTTCTGTTATGGACCTCCCTTGTGCTCTCCAGTGAA-----TAT 1000
Qy 259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db 1001 TCTATTGACAAAGTACTACGTCTCTTTCTGTTATTAAATGGGGCTTTTATGATTGGGA 1060
Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db 1061 CAGGCATCCCAAGCATTTGAAGCATTTGCAACCGCAAGAGAGAGCAGCTTATGAAATCTTC 1120
Qy 299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db 1121 AAGATAATTGCAATAAACCAGCATTTGACAGCTATTTCGAAGAGTGACATATAACACAGAT 1180
Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db 1181 AATATTAAAGGAAATTTGGAATTTCAAAATTTGTTCACTTACGTTACCTCTTCGAAAGAA 1240
Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
```

```
Db 1241 GTTAAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAAGATGGGCAGACAGTGGCGCTGGTT 1300
Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db 1301 GGGAAACAGTGGCTGCGGGAAGAGCACACGCTGCAGCTGATGCAGAGGCTCTATGACCCC 1360
Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db 1361 ACAGATGGCATGCTCTGTTATGATGCACAGGACATTTAGGACCATAATGTAAAGCATCTT 1420
Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db 1421 CGGAAATTAATCTGCTGCTGAGTCAGGAGCCTGTGTTGTTGCCACACGATAGCTGAA 1480
Qy 419 AsnIleLysTyrClyArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db 1481 AACATTCGCTATGCGCGCAAAATGTCACATGGATGAGATGAGAAAGCTGTTAGGAA 1540
Qy 439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db 1541 GCCAATGCCATGATTTTATCATGAACCTACCTAATAAATTTGACACTCTGTTGGAGAG 1600
Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db 1601 AGAGGGGCCAGCTGAGTGGTGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660
Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db 1661 CGAACCCCAAGATTTCTTCTGCGATGAGGCAACGTGAGCTCTGGACACTGGAAGTGAA 1720
Qy 499 SerAlaValGlnAlaLeuGluLys 507
Db 1721 GCAGTGGTTCAGTGGCCCTGGATAG 1747

RESULT 8
AAF86128
ID AAF86128 standard; cDNA; 4195 BP.
XX AAF86128;
AC AAF86128;
DT 25-JUN-2001 (first entry)
XX Cynomologous monkey P-glycoprotein cDNA variant 1.
DE Cynomologous monkey P-glycoprotein; PGP; multidrug transporter; MDRI;
KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRI;
XM efflux pump; 58.
XX Macaca fascicularis.
OS
XX Key Location/Qualifiers
CDS 100..3951
FT /*tag= a
FT /product= "PGP"
FT /note= "P-glycoprotein"
FT misc_feature 376..384
FT /*tag= b
FT /note= "Insertion of 9 nucleotides relative to PGP
FT allelic variant AAF86127"
XX
PN WO200123565-A1.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26592.
XX
XX 28-SEP-1999; 99US-0156921.
PR 12-OCT-1999; 99US-0158818.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX
```

DR WPI; 2001-316136/33.  
 XX P-PSDB; AAB81065.  
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 PT cell  
 XX  
 PS Example 1; Page 59-65; 84pp; English.  
 XX This invention relates to a polynucleotide sequence encoding a  
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDR1 is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention  
 CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents cDNA encoding  
 CC cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide  
 CC insert compared to the PGP allelic variant given in AAF86127.  
 XX  
 SQ Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1-71e-154 Length: 4195  
 Score: 1580.50 Matches: 306  
 Percent Similarity: 75.85% Conservative: 96  
 Best Local Similarity: 57.74% Mismatches: 103  
 Query Match: 60.98% Indels: 25  
 DB: 22 Gaps: 4  
 US-09-873-409-8 (1-514) x AAF86128 (1-4195)  
 QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 250 ATGGTGGTGGAACTTTGGCTGGCCATCCATCGAGCTGGAATTCCTCTCATGATGCTG 309  
 QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCys 33  
 Db 310 GTGTTTGGAGACATGACGGATACCTTTGCAAAATGACAGAAATTTAGGAGACT 369  
 QY 34 LeuValGlnThrAsnThrTyrrSer----- 41  
 Db 370 CTGTTGTTTAAACACACATATAGCAGTAATATCATCTATACAGTGCCTGCAATCTG 429  
 QY 42 -----PhePheArgLeuThrLeuTyrrValGlyIleGlyValAlaLeuIle 58  
 Db 430 GAGGAAGATATGACCGATGCTTATATATACAGTGGAAATTTGGTGGGTGCTGGTT 489  
 QY 59 PheGlyTyrrIleGlnIleSerLeuTrrIleIleThrAlaAlaArgGlnThrLysArgIle 78  
 Db 490 GCTGCTTACATTTACGTTTCATTTTGGTGGCTGGACCTGGAAAGACAAATACACAAAT 549  
 QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrrPheAspSerCysAsp 98  
 Db 550 AGAAACAGTTTTTTCATGCTATATACGACAGAGATAGCTGGTTTGTATGTCACGAT 609  
 QY 99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117  
 Db 610 GTTGGGAGCTTAAACACCGCTTACAGATGATGCTCTCCAAGATTAATGAAGGAATGGT 669  
 QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137  
 Db 670 GACAAAATTTGGAATGTCTTTCAGTCAATGCGACACATTTTTCACCTGGGTTTATAGTAGGA 729  
 QY 138 LeuValLysGlyTrrLysLeuLeuValThrLeuLeuValThrSerThrProLeuIleMetAla 157  
 Db 730 TTTACAGCTGTGGAGAGCTTAACCTTGTGATTTGGCCATTCAGCTCTCTTTGGAGCTG 789  
 QY 158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177

Db 790 TCAGTGCAGTCTGGCAAGATACTGTCTTATTACTGATAAAGAACTCTTAGCTTAT 849  
 QY 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197  
 Db 850 GCANAAGCTGGAGCAGTAGCTGAAGAGGCTTTGGCAGCAATTAGAACTGTGATGATTT 909  
 QY 198 ArgAlaGlnGluLysGluLeuGlnArgTyrrThrGlnAsnLeuLysAspAlaLysAspPhe 217  
 Db 910 GGAGGACAAAGAAAGAACTCGAAAGGTACACAAAATTTAGAAAGCTAAAGAAAT 969  
 QY 218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrrPhePheMetAsn 237  
 Db 970 GGGATAAAGAAAGCTATTACAGCCAAATATTTCTATAGGTGCTCTTCTCTGCTTATCTAT 1029  
 QY 238 GlyTyrrGlyLeuAlaPheTrrTyrrGlyTyrrSerLeuIleLeuAsnGlyGluProGly 257  
 Db 1030 GCATCTTATGCTCTGGCTTCTGGTATGGACCACTTGGTCTCTCAAGGAA----- 1083  
 QY 258 TyrrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrrCysIle 277  
 Db 1084 TATTTCTATTGGACAGTACTACTGTATTTCTTCTGTATTAATTGGGGCTTTTAGTGT 1143  
 QY 278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297  
 Db 1144 GGACGGCATCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGGACGCTTTTGAATC 1203  
 QY 298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrrLysPro 317  
 Db 1204 TTCAGATTAATGATTAATGAAGCCAGTATTTGACAGCTATTCGAAAGATGGGCAACCA 1263  
 QY 318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrrProSerArgPro 337  
 Db 1264 GATAATATTAAAGGAAATTTGGAATTCAGAAATTTTCACTTTCAGTTACCCATCTCGAAA 1323  
 QY 338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357  
 Db 1324 GAAGTTAAGATCTTGAAGGGCTGAACCTGAAGGTGCAGATGGGAGAGCGTGGGCCCTG 1383  
 QY 358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrrAsp 377  
 Db 1384 GTTGGAAACAGCGCTGTGGGAAGACACACACGCTCCAGCTGATGACAGAGCTTTATGAC 1443  
 QY 378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397  
 Db 1444 CCCACAGAGGGCATGCTCAGTGTGTGATGACAGGATATTAGGACCATAAACGTAAGGTT 1503  
 QY 398 TyrrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSer 417  
 Db 1504 CTACGGGAAATCATCGGTGTGTGAGTCAGGAACCTGTATTGTTTCCACACAGTAGCT 1563  
 QY 418 AsnAsnIleLysTyrrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArg 437  
 Db 1564 GAAACATTCGCTATGCTGCTGAAGATGTCACCATGATGAGATTCAGAAAGCTGCAAG 1623  
 QY 438 GluAlaAsnAlaTyrrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457  
 Db 1624 GAAGCCAAATGCTATGCTTATCATGAACTGCCTCAGAAATTTGCACACCTGGTTGA 1683  
 QY 458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477  
 Db 1684 GAGAGAGGGGCCAGCTGAGTGTGGTGGCAGAGACAGAGGATCGCCATTCACGTGCCCTG 1743  
 QY 478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497  
 Db 1744 GTTCGCAACCCCAAGATCTCTGCTGGACAGGCCACGTCAGCTTGGACACAGAAAGT 1803  
 QY 498 LysSerAlaValGlnAlaAlaLeuGluLys 507  
 Db 1804 GAAGCAGTGTTCAGGTGCTCTGGATAAG 1833  
 RESULT 9  
 AA249332



ID AAZ49332 standard; cDNA; 3860 BP.  
AC AAZ49332;  
XX  
XX 14-MAR-2000 (first entry)  
XX Human wild-type multidrug resistance-1 (MDR-1) cDNA.  
XX  
XX Multidrug resistance; MDR-1; P-glycoprotein;  
KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
KW gene therapy; gene replacement; genetic defect; thalassaemia;  
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
KW cytokine; wild-type; ds.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..3843  
XX FT /\*tag= a  
XX FT /product= "Human wild-type MDR-1 protein"  
XX FT replace (553..555, GTT)  
XX FT /\*tag= b  
XX FT /note= "cDNA sequence of G185V human mutant MDR-1 given  
XX in AAZ49333"  
XX  
XX WO9961589-A2.  
XX  
XX 02-DEC-1999.  
XX  
XX 27-MAY-1999; 99WO-US11825.  
XX  
XX 28-MAY-1998; 98US-0086988.  
XX  
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
XX Sorrentino B, Bunting K;  
XX  
XX WPI; 2000-072615/06.  
XX P-PSDB; AAY56186.  
XX  
XX Ex vivo expansion of haematopoietic stem cells transduced with a  
XX sequence encoding human multidrug resistance-1, used for bone marrow  
XX transplantation -  
XX  
XX Claim 10; Page 68-70; 113pp; English.  
XX  
XX This sequence represents cDNA encoding human wild-type  
XX multidrug resistance protein MDR-1. MDR-1 is a transmembrane  
XX efflux pump, responsible for the export of drugs from cells,  
XX particularly cancer cells. Wild-type MDR-1 shows increased  
XX resistance to etoposide and decreased resistance to vinca  
XX alkaloids compared with a mutant form (AAY58187) where the Gly at  
XX position 185 is replaced by Val. The invention relates to transducing  
XX haematopoietic stem cells with nucleic acid encoding an MDR protein  
XX and culturing the modified cells. The modified haematopoietic stem  
XX cells are useful in bone marrow transplantation (to reconstitute  
XX haematopoietic systems in patients who have undergone chemotherapy or  
XX radiation therapy) and in ex vivo gene therapy of genetic defects in  
XX cells derived from haematopoietic stem cells, e.g., thalassaemia,  
XX Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
XX cells can also be used to identify factors involved in regulating  
XX proliferation and differentiation in haematopoietic stem cells.  
XX Haematopoietic stem cells that express MDR-1 will be protected against  
XX chemotherapeutic agents, so can be engrafted while the patient is  
XX undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
XX provides sufficient cells to permit standard biochemical analysis.  
XX Overexpression of MDR-1 allows cytokine-driven expansion of  
XX haematopoietic stem cells by at least 10-fold compared with a maximum  
XX of 4-fold in known procedures.  
XX  
XX Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:  
Pred. No.: 2.18e-154 Length: 3860  
Score: 1579.00 Matches: 306  
Percent Similarity: 77.04% Conservative: 100  
Best Local Similarity: 58.08% Mismatches: 99  
Query Match: 60.92% Indels: 22  
DB: 21 Gaps: 5  
US-09-873-409-8 (1-514) x AAZ49332 (1-3860)  
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
DB 151 ATGGTGGTGGGAACCTTGGCTGCCATCCATCCATGGGGCTGGACTTCCTCATGATGCTG 210  
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
DB 211 GTGTTGGAGAAATGACAGATATCTTGGCAATGACAGAAATTTAGAAGATCTGATGTCA 270  
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
DB 271 AACATCACTAATAGAAAGTGATATCAATGATACAGGGCTTCTTCATGAATCTGGAGGAAGAC 330  
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
DB 331 ATGACCAGATATGCCCTATTATTACAGTGGAAATTTGGTGGGGTGGCTGCTGCTTAC 390  
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
DB 391 ATTCAAGTTTCATTTGGTGGCTGGCAGCTGGAGAGATAGGCTGTTGATGTGCACGATGTGGGAG 450  
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
DB 451 TTTTTCATGCTATATAGTCACAGAGATAGGCTGTTGATGTGCACGATGTGGGAG 510  
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
DB 511 CTTAACACCCGACTTACAGATGATGTCTCTAGATTAATAGAGGTATTGGTGACAAAATT 570  
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
DB 571 GGAATGTCTTCAGTCAATGGCAACATTTTTCACCTGGGTATTATAGTAGGATTACACGT 630  
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
DB 631 GGTGGGAAGCTAACCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTGAGTGTCT 690  
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
DB 691 GTCTGGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750  
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
DB 751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACCTGTGATTGCAATTGGAGACAA 810  
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
DB 811 AAGAAGAACTTGAAGGTACACAAAATTTAGAAAGAACTAAAGAAATGGGATAAAG 870  
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
DB 871 AAAGCTATTACAGCCAATTTCTATAGGTGCTGCTTCTCTGCTGATGATGATCTTAT 930  
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
DB 931 GCTCTGGCTTCCTGGATGGGACACCTTGGTCTCTCAGGGGAA-----TATTCTATT 984  
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
DB 985 GGACAAGTACTACTGTATTCCTTTCTGATTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1044  
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
DB 1045 TCTCCAAGCATTTGAAGCAATTTGCAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 1104







CC of mdrlb2 activity. It may also be used to establish assays to  
 CC predict oral absorption and pharmacokinetics of drugs in humans, and  
 CC thus enhance the design of formulations through the identification of  
 CC compounds with optimal development characteristics (i.e., high oral  
 CC bioavailability, UID (once a day) dosing, reduced drug interactions,  
 CC reduced variability, and reduced food effects), specifically to  
 CC avoid interactions with human mdrlb2. Transgenic and knockout animals  
 CC created using DNA encoding the rat mdrlb2 may be used to gain an  
 CC insight into treating and preventing human diseases such as cancer,  
 CC inflammation, cardiovascular disease, central nervous system disorders,  
 CC autoimmune disorders and kidney disease. The present sequence represents  
 CC cDNA encoding rat mdrlb2.  
 XX  
 SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:  
 Pred. No.: 5.12e-154 Length: 4233  
 Score: 1576.00 Matches: 308  
 Percent Similarity: 75.90% Conservative: 92  
 Best Local Similarity: 58.44% Mismatches: 105  
 Query Match: 60.80% Indels: 22  
 DB: 22 Gaps: 4

US-09-873-409-8 (1-514) x AAF27498 (1-4233)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
 Db 170 ATGGCTCTGGGAACCTCGCTGCTATCATCCAGCAACCTCGCTCCCTCTGATGCTG 229  
 Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeu----- 30  
 Db 230 GTGTTGGATACATGACAGATAGTATTTTACCAAGCAGAGACCCGCAATTCGCCGAGCGTT 289  
 Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45  
 Db 290 ACTAATCAAAAGTGAATCAACAGTACACAGACCGTCAGCGACAGCAGCTCTGGAGGAGGAC 349  
 Qy 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaLeuIlePheGlyTyr 61  
 Db 350 ATGGCCATGACCGCTACTATTACCGGCATTTGGTCGGTGTGCTCATCGTTGGCTAC 409  
 Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrIleArgLysGln 81  
 Db 410 ATCCAGGTTTCACITTTGGTGGCTGGCAGCTGGGAGACAAATACACAAGATTAGGCACAAG 469  
 Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
 Db 470 TTTTTCATGCCATCATGAATCAGGAGATAGGCTGGTTTGCACGTGAATGACGCTGGGGAG 529  
 Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
 Db 530 CTCACACCCGGCTCACAGATGAGCTCTCCAAATTAATGACGGAATTTGGTACAAACTT 589  
 Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
 Db 590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTATATAGGATTTATAGT 649  
 Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
 Db 650 GGTGTGAAGCTTAACCTTGTAATTTTGGCCGCTCAGCCCTCTTATTGGGTGTGTCATCTGCC 709  
 Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
 Db 710 ATGTGGGCAAGGACTACTCTTACTTAATAGGAACCTCCAGGCTTATGCCAAGAGCT 769  
 Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
 Db 770 GGACAGTTTGGCGAAGAAGTCTTAGCACCATCAGACTGTGATGGTTTGGAGGACAA 829  
 Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
 Db 830 AAGAAGGAACCTTGAAGGTCAATAAATAATTTAGAAAGAGCTAAAGAGAGTTGGCATAAAG 889

Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
 Db 890 AAAGCCATCAGCGCCCAACATTTCCATAGGTATTCCTGTTGCTTATGCTGCTTAT 949  
 Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260  
 Db 950 GCATCGGCATTCGTGTATGGGACCTCTTGGTCTCTCAAAATGAA-----TATTCTATT 1003  
 Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerTyrCysIleGlyAlaAla 280  
 Db 1004 GGCAAGTCTTACCGTCTCTCTCTATTATTATGGGACCTTTCAGTATTGGACATTTA 1063  
 Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
 Db 1064 GCCCAACACATAGAGCCTTTGCAATGCAAGAGGGGCGAGCCTATGAAATCTTCAAGATA 1123  
 Qy 301 IleAspLysIlePheProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
 Db 1124 ATTGATAATGAGCCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACGACAGATATA 1183  
 Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
 Db 1184 ATGGGAAATTTGGAATTTAAATGTTTACTTCACTACCATCACAAAGTGAAGTTAAG 1243  
 Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
 Db 1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGCGACAGCGTAGCCCTGGTTGGCAAC 1303  
 Qy 361 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
 Db 1304 AGTGGTGTGGGAAAGCAACCTGTCCAGCTCTGCAGAGGCTCTACGACCCCATAGAG 1363  
 Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
 Db 1364 GCGCAGCTCAGTATCCAGGACAGGACATCAGGACCATCAATGTGAGGTATCTCGCGGAA 1423  
 Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
 Db 1424 ATCATTGGGTGTGAGTCAGGAACCCGCTGTTTGGCCACCAGCATTCGCGAACAACATT 1483  
 Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
 Db 1484 CGTATGGCGGAAACCGTCACCATGTGATGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543  
 Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460  
 Db 1544 GCCTATGACTTCAATGAAATCTGCCCCCAAAATTTAACCCCTGGTGTGTGAGAGGG 1603  
 Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
 Db 1604 GCGCAGCTGAGTGGGGACAGAAACAGAGGATCGCATTTGCCCGGCGCTTGTCCGCAAC 1663  
 Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
 Db 1664 CCCAAGATCCTTTTGTGTGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGGCGTG 1723  
 Qy 501 ValGlnAlaAlaLeuGluLys 507  
 Db 1724 GTTCAGCGCCCTCTGGATAG 1744  
 RESULT 13  
 AAD03488  
 ID AAD03488 standard; cDNA; 4279 BP.  
 XX  
 AC AAD03488;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Dog P-glycoprotein (PGP) cDNA #1.  
 XX  
 KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;  
 KW drug bioavailability; transgenic animal; genetic model; ss.  
 XX







Db 997 TCATATGCTCTGGCTTCTGGTATGGACCTCTTGCTCTCCAGTGA-----TAT 1050

Qy 259 ThrIleGlyThrValLeuAlaValPheSerValIleHisSerSerTyrcysIleGly 278

Db 1051 ACTATTGGACAGGTACTCACTGCTCTCTTCTGTTATTAATTGGGGCTTTAGTATTGA 1110

Qy 279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePhe 298

Db 1111 CAGGCATCCCAAGCATTTGAAGCATTTGCAAGCGCAAGAGGAGCAGCTTATGAATCTTC 1170

Qy 299 GlnValIleAspLysPheProSerIleAspAsnPheSerThrAlaGlyTyrcysProGlu 318

Db 1171 AAGATATTGCAATTAACCAAGCATTTGAGAGTATTCGAAGAGTGCATAAACAGAT 1230

Qy 319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrcysProSerArgProSer 338

Db 1231 AATATTAAAGGGAATTTGGAATTCAAATATGTTCACTTTCAGTTACCTTCTCGAAAGAA 1290

Qy 339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358

Db 1291 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTTCAGAGTGGCGAGCAGTGGCGCTGTT 1350

Qy 359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrcysPro 378

Db 1351 GGGAAAGTGGCTGCGGAAAGGACGACCGTGCAGCTGATGAGAGGCTTATGACCCC 1410

Qy 379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398

Db 1411 ACAGATGGCATGCTCTGTTATGATGGACAGGACATTAGGACCATTAATGTAAAGCATCTT 1470

Qy 399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418

Db 1471 CGGAAATTAAGTCTGCTGGTGTGAGTCAAGGCTGTGTGTTGTTGCCACCATAGCTGAA 1530

Qy 419 AsnIleLysTyrcysValArgAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438

Db 1531 AACATTCGCTATGCGCGGAAATATGACCATGATGAGATGAGATGAGAAAGCTTTAAGGAA 1590

Qy 439 AlaAsnAlaTyrcysPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458

Db 1591 GCCAATGCCTATGATTTTATCATGAACACTACCTAATAAATTTGCACACTCTGGTTGGAGAG 1650

Qy 459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478

Db 1651 AGAGGGCGCGCTGAGTGTGTGGACAGAAACAGAGAATCGCCATGCTCGGGCCCTGGTT 1710

Qy 479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498

Db 1711 CGCAACCCCAAGATTTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGGAAGTGAA 1770

Qy 499 SerAlaValGlnAlaAlaLeuGluLys 507

Db 1771 GCAGTGTTCAGTGGCCCTGGATAG 1797

RESULT 15

AAV32645

ID AAV32645 standard; cDNA; 4669 BP.

XX AAV32645;

AC AAV32645;

XX AAV32645;

DT 23-SEP-1998 (first entry)

XX Human P glycoprotein (Pgp) cDNA.

DE Human P glycoprotein; Pgp; multi-drug resistance; cancer;

KW UIC2 monoclonal antibody; mAB; cytotoxic; transmembrane efflux pump; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FH 1..424

FT 5'UTR

FT 425..4267

FT CDS

FT

FT /\*tag= b

FT /product= "Wild-type Pgp"

FT 4268..4669

FT /\*tag= c

XX 3'UTR

XX WO9821325-A1.

XX 22-MAY-1998.

XX 17-NOV-1997; 97WO-US21214.

XX 15-NOV-1996; 96US-0752447.

XX (INGE-) INGENEX INC.

PA (UNITI) UNIV ILLINOIS FOUND.

XX Mechetner E, Roninson IB;

XX WPI; 1998-297930/26.

DR P-PSDB; AAW48997.

XX Immunological reagent specific for P-glyco:protein - useful for

PT detecting multi-drug resistant cancer, isolating haematopoietic

PT cells and selective cell killing

XX Claim 1; Fig 1A; 89pp; English.

XX The present sequence represents the human P glycoprotein (Pgp)

CC cDNA which encodes the wild-type Pgp protein. Pgp is a transmembrane

CC efflux pump protein involved in multi-drug resistance of cancer cells.

CC The invention provides methods for developing and using immunological

CC reagents specific for certain mutant forms of Pgp and wild-type Pgp in

CC a conformation associated with substrate binding or in the presence

CC of ATP depleting agents. An example of the immunological reagent

CC is the UIC2 monoclonal antibody (mAB). mAB UIC2 specifically binds

CC to Pgp in a particular biochemical conformation and is capable of

CC inhibiting drug efflux from Pgp-expressing cells. The immunological

CC reagents are claimed to be useful for detecting Pgp expression in

CC mammalian cells, including low level expression, particularly in cancer

CC cells to diagnose multi-drug resistance. The invention claims that

CC these immunological reagents are more specific than known reagents for

CC detecting Pgp and they also eliminate the need for costly and laborious

CC screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.

XX Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 5.9e-154 Length: 4669

Score: 1576.00 Matches: 305

Percent Similarity: 77.04% Conservative: 101

Best Local Similarity: 57.87% Mismatches: 99

Query Match: 60.80% Indels: 22

DB: 19 Gaps: 5

US-09-873-409-8 (1-514) x AAV32645 (1-4669)

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20

Db 575 ATGGTGGTGGAACTTTGGCTGCCATCATCCATCGGGCTGGAGCTTCTCTCATGATGCTG 634

Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31

Db 635 GTGTTTGGAGAAATGACAGATATCTTTGCCAATGCAGGAAATTTAGACAGATCTGATGTC 694

Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrcysPhePhe----- 43

Db 695 AACATCCTAATAAGATGATATCATGATGATGATGATGATGATGATGATGATGATGATG 754

Qy 44 -----ArgLeuThrLeuTyrcysValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 755 ATGACCAGGTATGCCCTATTATTACAGTGGAAATTTGGGGTGTGCTGCTGCTTAC 814

Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81

```
|||||:||||| ||| :|| ||||| :||||| |||||
815 ATTCAGGTTTCATTTGGTGCTGGCAGCTGGAGACAAAATACACAAAATTAGACAAACAG 874
QY PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
875 TTTTTCATGCTATTATGCGACAGAGATAGGCTGGTTTGTATGTCACGATGTTGGGGAG 934
QY LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
935 CTTAACACCCGACTTACAGATGATGCTCCAGATTAATGAAGGAATGGTGACAAATTT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
995 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1055 GGTGGGAAGCTTAACCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACGTGTCAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1115 GTCTGGGCAAGATACCTATCTTCACTTACTGATPAAAGAACTCTTAGCGTATGCCAAGCT 1174
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1175 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATTGCAATTTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1235 AAGAAAGAACTTGAAGAGGTCAACAAAATTTAGAAAGCTAAAGAAATTTGGGATAAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1295 AAAGCTATTACAGCCCAATATTTCTATAGTGTCTTCTCTGCTGATCTATGCACTTAT 1354
QY 241 GlyLeuAlaPheThrTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1355 GCTGTGGCCTTCTGGTATGGACCACCCTTGGTCTCTCAGGGGAA-----TATCTTATT 1408
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1409 GGACAAGTACTCCTGTTATCTTCTTCTGTATTAATTTGGGGCTTTTAGTGTGGACAGGCA 1468
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1469 TCTCCAGCATTGAAGCATTTCGCAATGCAAGAGCAGCAGCTTATGAAATCTTCAAGATA 1528
QY 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1529 ATTGATAATAAGCCCAAGTATTGACAGCTATTGCAAGAGTGGGCAACAACCCAGATAATTT 1588
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1589 AAGGGAAATTTGGAAATTCAGAAATGTTTCACTTCAGTTACCCATCTCGAAAAGAGTTAAG 1648
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1649 ATCTTGAGGGCCTGAACCTGAAGCTGCAGAGTGGGAGCGGTGGCCCTGTTGGAAAC 1708
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1709 AGTGGCTGTGGGAAGAGACACAACAGTCAGCTGATGTCAGAGGCTCTATGACCCCCACAGAG 1768
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1769 GGGATGCTGAGTGTGTGAGTGGACAGGATATTAGGACCATAATGTAAGGTTTCTACGGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1829 ATCATTTGGTGTGGTGTGAGTCAGGAACCTGTATTGTTGCCACCAGATAGCTGAAACATT 1888
QY 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 1889 CGCTATGGCCGTGAAAATGTCCACCATGGATGAGATTGAGAAAAGCTGTCAAGGAAGCCAAT 1948
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1949 GCTATGACTTTATCATGAACTGCTCATATAATTTGACACCTTGGTTGGAGAGAGGGG 2008
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2009 GCCCAGTTGAGTGGTGGGAGAGAGAGGATCGCCATTGCACGTGCCCTGGTTTCGCAAC 2068
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaIleAspSerGluSerLysSerAla 500
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2069 CCCAAGATCCTCTCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAAGCGAGCAGTG 2128
QY 501 ValGlnAlaAlaLeuGluLys 507
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2129 GTTCAGGTGGCTCTGGATAG 2149
```

Search completed: March 30, 2003, 03:22:28  
Job time : 555.349 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 03:22:44 ; Search time 3535.38 Seconds  
(without alignments)  
2354.618 Million cell updates/sec

Title: US-09-873-409-8  
Perfect score: 2592  
Sequence: 1 MILGILASLVNGACILPLMLP.....SESKSAVOALEKDTPTVSP 514

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh

-O=/cgn2\_1/USPTO spooll/US9873409/runat\_27032003.115420.19240/app\_query.fasta\_1.7544  
-DB=EST -QFMT=fastp -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US9873409 @CGN 1.1.30544 @runat\_27032003.115420.19240 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSEEXT=7

Database :

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: gb\_gas.\*
- 18: em\_gas\_hum.\*
- 19: em\_gas\_inv.\*
- 20: em\_gas\_pln.\*
- 21: em\_gas\_vrt.\*
- 22: em\_gas\_fun.\*
- 23: em\_gas\_nam.\*
- 24: em\_gas\_mus.\*
- 25: em\_gas\_other.\*
- 26: em\_gas\_pro.\*
- 27: em\_gas\_rtd.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092	42.1	669	13	BM013981	BM013981 603639619
2	1068.5	41.2	1033	14	BM926413	BM926413 AGENCOURT
3	1029	39.7	760	13	BM016204	BM016204 603642659
4	926.5	35.7	929	14	BQ882401	BQ882401 AGENCOURT
5	918	33.4	545	9	AL040762	AL040762 DKF2P434C
6	856.5	33.0	948	13	BM471690	BM471690 AGENCOURT
7	750	28.9	894	12	BF584668	BF584668 602098406
8	744.5	28.7	872	11	AK020318	AK020318 Mus muscu
9	710	27.4	2676	11	AK014319	AK014319 Mus muscu
10	671.5	25.9	803	12	BQ298756	BQ298756 602396681
11	659	25.4	943	9	ALJ520322	ALJ520322 ALJ520322
12	654	25.2	405	9	AA243820	AA243820 zr67906.r
13	650	25.1	550	13	BJ075336	BJ075336 BJ075336
14	637	24.6	780	12	BG587938	BG587938 EST489713
15	625	24.1	559	12	BF692596	BF692596 602248949
16	621	24.0	606	14	BQ262340	BQ262340 f285a08.x
17	621	24.0	681	14	BU006064	BU006064 QGGSN16.Y
18	604.5	23.3	932	17	AZ670821	AZ670821 ENTJN69TF
19	602.5	23.2	913	17	BH155700	BH155700 ENTRO54TR
20	599	23.1	880	17	AZ687805	AZ687805 ENTJUS2TF
21	596.5	23.0	871	17	AZ682350	AZ682350 ENTXB18TF
22	596.5	23.0	2564	11	AK020022	AK020022 Mus muscu
23	590.5	22.8	891	17	AZ682250	AZ682250 ENTKT68TR
24	585.5	22.6	853	17	AZ679807	AZ679807 ENTHI67TR
25	584.5	22.6	1813	11	U66688	U66688 Homo sapien
26	582.5	22.5	787	12	BG584063	BG584063 EST485823
27	581	22.4	586	13	BQ336461	BQ336461 BU336461
28	579	22.3	661	12	BG303675	BG303675 f136b04.x
29	578	22.3	575	10	AW342644	AW342644 f181f06.x
30	577.5	22.3	886	17	AZ540627	AZ540627 ENTQ18TF
31	577	22.3	835	13	BJ349604	BJ349604 BJ349604
32	576.5	22.2	947	17	AZ683753	AZ683753 ENTJL96TF
33	576	22.2	547	13	BQ333166	BQ333166 BJ333166
34	574.5	22.2	885	17	AZ686798	AZ686798 ENTF228TR
35	574	22.1	542	13	BQ339421	BQ339421 BJ339421
36	572.5	22.1	860	17	AZ543293	AZ543293 ENTQ80TF
37	572.5	22.1	886	17	BH139685	BH139685 ENTNA47TR
38	571	22.0	715	14	BQ869512	BQ869512 QGD6G09.Y
39	570	22.0	814	13	BQ352304	BQ352304 BJ352304
40	570	22.0	820	13	BQ356540	BQ356540 BJ356540
41	570	22.0	1019	12	BG248052	BG248052 602359987
42	569	22.0	1341	11	AY108285	AY108285 Zea mays
43	565	21.8	657	14	BQ802058	BQ802058 WHE2821.F
44	564	21.8	636	13	BQ386538	BQ386538 BJ386538
45	564	21.8	750	13	BU445883	BU445883 BJ445883

ALIGNMENTS

RESULT 1  
BM013981  
LOCUS 603639619F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:5415556 5', linear EST 30-OCT-2001  
DEFINITION mRNA sequence.  
ACCESSION BM013981  
VERSION BM013981.1 GI:16528335  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 669)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12060 row: f column: 05  
High quality sequence stop: 664.  
Location/Qualifiers  
1..669  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NIH MGC 87"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH MGC Library."  
BASE COUNT 203 a 122 c 162 g 182 t

## ORIGIN

Alignment Scores:  
Pred. No.: 1,77e-119 Length: 669  
Score: 1092.00 Matches: 215  
Percent Similarity: 99.54% Conservative: 2  
Best Local Similarity: 98.62% Mismatches: 0  
Query Match: 42.13% Indels: 1  
DB: 13 Gaps: 0

US-09-873-409-8 (1-514) x BM013981 (1-669)

QY 265 AlavalPheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 284  
Db 17 AGTGTTTCTTAATGTAATCCATAGCAGTATTGCATTGGAGCAGAGTCCCTCACTTT 76  
QY 285 GluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnValIleAspLysLys 304  
Db 77 GAACCTTCGCAATAGCCGAGGAGTCCCTTCATATTTTCCAGGTATTGATAAGAAA 136  
QY 305 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 324  
Db 137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGAACTGTG 196  
QY 325 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 344  
Db 197 GAATTTAAAATGTTCTTTCAATATATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256  
QY 345 LeuAsnLeuArgIleLysSerGlyValAlaLeuValGlyLeuAsnGlySerGly 364  
Db 257 CTGAATCTCAGAAATAGTCTGGAGACACAGTCGCTTGGTCTCATAGCAGTGGG 316  
QY 365 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 384  
Db 317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTATATCATCGGATGATGCTTTATCATG 376  
QY 385 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 404  
Db 377 GTGGATGAGAATGCAATCAGAGCTTTTAAATGTGGCGGATTTATCGAGACCATATTGGAGTG 436  
QY 405 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAniIleLysTyrGlyArg 424  
Db 437 GTTAGTCAAGAGCCTGTTTTGTCGGGACACCATCATGTAACATATCAAGTATGACGA 496  
QY 425 AspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 444  
Db 497 GATGATGTGACTGATGAAGATGAGAGAGCAGCAAGGGAAGCAAAATGCGTATGATTTT 556

QY 445 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGlyAlaGlnMetSer 464  
Db 557 ATCATGGAGTTTCTTAATAATTAATATCATTTGGTAGGGAAAAAGAGCTCAATGAGT 616  
QY 465 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 481  
Db 617 GGAGGGCAGAAAACAGAGGCATCGCAATTGCTGCTTAGTTTGGAAACCCC 668

## RESULT 2

BM926413  
LOCUS  
DEFINITION 5', mRNA sequence.  
AGENCOURT 6600788 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5764845  
BM926413  
VERSION BM926413.1 GI:19376792  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1033)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12818 row: o column: 22  
High quality sequence stop: 3  
High quality sequence stop: 686.

## JOURNAL

## COMMENT

## FEATURES

## source

1..1033  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5764845"  
/clone\_lib="NIH MGC\_114"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
male brains, age range 23-27 yo. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.5 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 019. Note:  
this is a NIH MGC Library."  
BASE COUNT 289 a 212 c 218 g 312 t

## ORIGIN

## Alignment Scores:

Pred. No.: 2.3e-116 Length: 1033  
Score: 1068.50 Matches: 229  
Percent Similarity: 93.17% Conservative: 3  
Best Local Similarity: 91.97% Mismatches: 17  
Query Match: 41.22% Indels: 4  
DB: 14 Gaps: 0

US-09-873-409-8 (1-514) x BM926413 (1-1033)

QY 266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
Db 180 GTTTTCTTAGTGATATCCATAGCAGTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 239  
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305  
Db 240 ACCTTCGAATAGCCCGAGGAGTGCCTTTTCATATTTTCCAGGTCTATTGATAAGAAACC 299

QY	306	SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu	325
Db	300	AGTATAGATAAATTTTCCACAGCTGGATATAAACCCTGAATCCATAGAAAGGAACTGTGAA	359
QY	326	PheIleAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu	345
Db	360	TTTAAATAATGTTTCTTCAATATATCCATCAAGACCATCTATCAAGATTCGAAAGGTCG	419
QY	346	AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys	365
Db	420	AATCTCAGATTAAATCTGGAGACAGTCGCTTGGTCGGTCTCAATGCGAGTGGGAAG	479
QY	366	SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal	385
Db	480	AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGCTTTATCATGGTG	539
QY	386	AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal	405
Db	540	GATGAGATGACATCAGAGCTTTAAATGTGGCGATTTATCGAGACCATATGGAGTGGTT	599
QY	406	SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp	425
Db	600	AGTCAAGAGCTGTTTGTTCGGGACCAACCATCAGTAACATATCAAGTATGACGAGAT	659
QY	426	AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle	445
Db	660	GATGTGACTGATGAAGAGATGGAGAGACGACGAAGGGAAGCAATGCGTATGATTTATC	719
QY	446	MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly	465
Db	720	ATGGAGTTTCTCTAATAA-TTTAATACATTTGGTAGGGGAANAAGGAGCTCCAAATCAGTGA	778
QY	466	GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle	485
Db	779	AGCCAGAACCG-AGGATCCCAATTCCTTCCTGCTTAGTTTGGAAACCCCGAT-CTGAAT	836
QY	486	LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeu	505
Db	837	TTAGATGAAGCTAGCTGCTCCCTGGATTCAGAAACCAAGTCAACCTGGTCA--GCTGGGCTG	894
QY	506	GluLysAspThrProArgTyrSerPhe	514
Db	895	GAGAAGGAAACCCCGGNAATTCTTT	921
RESULT	3		
LOCUS	BM016204	760 bp mRNA linear EST 30-OCT-2001	
DEFINITION	603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',		
ACCESSION	BM016204	mRNA sequence.	
VERSION	BM016204.1	GI:16530558	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 760)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-r@mail.nih.gov		
	Tissue procurement: DCTD/DTF		
	CDNA Library Preparation: Life Technologies, Inc.		
	CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: L1AM12068 row: e column: 16		
	High quality sequence stop: 740.		
	Location/Qualifiers		

FEATURES

source	1..760		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:5418615"			
/clone_lib="NIH_MGC_87"			
/tissue_type="mammary adenocarcinoma, cell line"			
/lab_host="DH10B (phage-resistant)"			
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;			
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.			
Average insert size 1.383 kb. Library enriched for			
full-length clones and constructed by Life Technologies.			
Note: this is a NIH_MGC Library."			
BASE COUNT	229 a 150 c 187 g 193 t		
ORIGIN			
Alignment Scores:	7.18e-112 Length: 760		
Pred. No.:	1029.00 Matches: 209		
Score:	98.12% Conservatives: 0		
Percent Similarity:	98.12% Mismatches: 2		
Best Local Similarity:	39.70% Indels: 2		
Query Match:	13 Gaps: 0		
DB:			
US-09-873-409-8 (1-514) x BM016204 (1-760)			
QY	303	LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly	322
Db	3	AGAAACCCAGTATAGATAAATTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGA	62
QY	323	ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu	342
Db	63	ACTGTGGAATTTTAAATAATGTTTCTTCAATATATCCATCAAGACCATCTATCAAGATTCG	122
QY	343	LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly	362
Db	123	AAAGTCTGAATCTCAGATTAACTCTGGAGACAGTGCCTTGGTTCGATCAATGGC	182
QY	363	SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe	382
Db	183	AGTGGGAAGAGTAGCTAGTCCAGCTTCTCGAGAGCTTATATGATCCGGATGATGCTTT	242
QY	383	IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle	402
Db	243	ATCATGCTGGATGAGAATGACATCAGAGCTTTAAATGTGGCGCATTTATCGAGACCATATN	302
QY	403	GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr	422
Db	303	GGAGTGGTTAGTCAAGAGCTGTTTGTTCGGGACCAACCATCAGTAACTATCAAGTAT	362
QY	423	GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr	442
Db	363	GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAATGCGTAT	422
QY	443	AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGln	462
Db	423	GATTTTATCATGAGTTCCTTAATAATTTAATACATTCGTAGGGGAAAAGAGCTCAA	482
QY	463	MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLys	482
Db	483	ATGAGTGGAGGGCAGAAACAGAGATCGCAATTTGGCTTCGCTTTCGAAACCCCAA	542
QY	482	stleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG	502
Db	543	GATTCGATTTTCAGATGAGCTACGCTCGCTCGCTGGATTCAGAAAGCAAGTCAGCTGTC	602
QY	502	lnAlaAlaLeuGluLysAspThrProArgTyrSer	513
Db	603	AAAGTGCACCTGGAGAAGGATACCCCGAGGTATTCC	637
RESULT	4		
LOCUS	BQ882401	929 bp mRNA linear EST 16-AUG-2002	
DEFINITION	AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782		

```

5', mRNA sequence.
BO882401
VERSION
BO882401.1 GI:22274409
EST.
SOURCE
HOMO
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 929)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2492 row: c column: 15
High quality sequence stop: 677.
FEATURES
source
1..929
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6291782"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC Library."
BASE COUNT 281 a 176 c 237 g 234 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1..71e-99 Length: 929
Score: 926.50 Matches: 188
Percent Similarity: 79.30% Conservatives: 61
Best Local Similarity: 59.87% Mismatches: 58
Query Match: 35.74% Indels: 7
DB: 14 Gaps: 2
US-09-873-409-8 (1-514) x BO882401 (1-929)
QY 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleAArgThrValII 195
Db 1 GCGTATGCAAAAGCTGAGCAGTACCTGAGAGAGCTTGGCAGCAATTAGACCTGGTAT 60
QY 195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaIy 215
Db 61 TGCATTTGGAGGACAAAGAAAGAACTTGAAGAGGTACAACAAAATTTAGAAAGCTAA 120
QY 215 sAspPheGlyIleLysArgThrIleAlaSerLysValSerLysAlaValTyrPhePh 235
Db 121 AAGNATTTGGGATAAAGAGCTATTACAGCAATATTTCTATAGGTCTGCTTTCCTGCT 180
QY 235 eMetAsnGlyTyrThrGlyLeuAlaPheTyrThrGlyThrSerLeuIleLeuAsnGlyGI 255
Db 181 GATCTATGCATCTTATGCTCTGGCTCTTGGTATGGGACCACTTGGCTCTCTCAGGGGA 240
QY 255 uProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSery 275
Db 241 A-----TATTCATTAGGACCAAGTACTACGTATTTCTGTATTAATTTGGGGCTTT 294
QY 275 rCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPh 295

```

```

Db 295 TAGTGTGGACAGGCATCTCCAAAGCATTTGAAGCATTTGCAAAATGCAAGAGGACCTTA 354
QY 295 eHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTY 315
Db 355 TGAATCTCTCAGATTAATTTGATAATAGCCAAATATTGACGCTATTGCAAGAGTGGGCA 414
QY 315 rLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSe 335
Db 415 CAAACCCAGATATATTAGGGAAATTTGGAATTCAGAAATGTTCACTTCACTTACCCATC 474
QY 335 rArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVa 355
Db 475 TCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTTGAAGGTGCAGAGTGGCAGACGGT 534
QY 355 lAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLe 375
Db 535 GGGCCCTGGTGGAAACAGTGGCTGTGGGAAGACACAAACAGTCCAGCTGATGAGAGGCT 594
QY 375 uTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVa 395
Db 595 CTATGACCCCAACAGAGGGATGGTCAGTGTGTGATGGACAGGATATTAGCACCAATAATGT 654
QY 395 lArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrTh 415
Db 655 AAGGTTTCTACGGGAAATCATTTGGTGTGAGTACAGGAACCTGTATTGTTGCCACCAC 714
QY 415 rIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgal 435
Db 715 GATAGCTGAAAACATTTCCCTATGGCGGTGAAATGTCACCATGATGAGATGAGATGAGAAAGC 774
QY 435 aAlaArgGluAlaAsnAla-TyrAspPheIleMetGluPheProAsnLysPheAsnThrL 455
Db 775 TGTCACAGGAGCAATGCCCTATGACTTTATCATGAACTGCTCATATAATTTGACACCC 834
QY 455 euValGlyGluLysGlyAlaGlnMetSerGlyGly---GlnLysGlnArgIleAlaIleA 474
Db 835 TGGTTGGANAAGAGAGGGCCAAATTCAGTGGTGGGCGACAAACCAAGGATCTCCCTTGG 894
QY 474 lAArg-AlaLeuVal-ArgAsnProLysIleLeu 484
Db 895 CACGTGGCCCTGGTTCCTCCCAACCCCAAGATCTTC 928
RESULT 5
AL040762 545 bp mRNA linear EST 29-FEB-2000
LOCUS
DEFINITION
DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFZp434C1815 5', mRNA sequence.
ACCESSION
AL040762
VERSION
AL040762.1 GI:5409708
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 545)
BLOECKER, H., BOECKER, M., BRANDT, P., MEWES, H.W., GASENHUBER, J. and
WIEMANN, S.
EST (Bioecker, et al.)
Unpublished (1999)
Contact: Bioecker H
MIPS
Am Klopferstr. 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
sl sequence also available.
This clone (DKFZp434C1815) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

```

FEATURES
source      Location/Qualifiers
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p434C1815"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH108"
/notes="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT  171 a  89 c 138 g 147 t
ORIGIN

Alignment Scores:
Pred. No.:      7,34e-99      Length:      545
Score:          918.00      Matches:    181
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    35.42%      Indels:     0
DB:              9      Gaps:      0

US-09-873-409-8 (1-514) x AL040762 (1-545)

QY 306 SerIleAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 325
|
|
|
Db 3 AGTATAGATAAATTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTGGAA 62
|
|
|
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
|
|
|
Db 63 TTTAAATAATGTTCTTCAATATTCATCAAGACCATCTATCAAGATCTCGAAGGTCGT 122
|
|
|
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
|
|
|
Db 123 AATCTCAGAATTAAGTCTGGAGACAGTCGCTTGGTCGCTCTCAATGGCAGTGGGAAG 182
|
|
|
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
|
|
|
Db 183 AGTACGGTAGTCCAGCTCTCGACAGAGTTATATGATCCGATGATGCTTTATCATGGTG 242
|
|
|
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
|
|
|
Db 243 GATGAGAAATGACATCAGAGCTTTAAATGTGGGCAATTTATCGAGACCATATGGATGGTT 302
|
|
|
QY 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
|
|
|
Db 303 AGTCAAGAGCTGTTTGTTCGGGACCAACCATCAGTAACAATATCAAGTATGACAGAGAT 362
|
|
|
QY 426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
|
|
|
Db 363 GATGTGACTGATGAAGAGATGGAGAGACAGCAAGGGAAGCAATGCGTATGATTTTATC 422
|
|
|
QY 446 MetGluPheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGly 465
|
|
|
Db 423 ATGAGGTTTCTTAATAAATTTAATACATTTGGTAGGGAAAAGAGCTCAAAATGAGTGA 482
|
|
|
QY 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
|
|
|
Db 483 GGGCAGAAACAGAGATCGCAATTTGCTCGTCTTAGTTTCGAACCCCAAGATTCGATT 542
|
|
|
QY 486 Leu 486
|
|
|
Db 543 TTA 545

RESULT 6
LOCUS      BM471690
DEFINITION AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cdna clone IMAGE:5539117
5', mRNA sequence.
ACCESSION  BM471690
VERSION    BM471690.1 GI:18520732
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 948)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM12233 row: b column: 14  
High quality sequence stop: 569.  
FEATURES  
source Location/Qualifiers  
1..948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5539117"  
/clone\_lib="NIH MGC 72"  
/tissue\_type="melanoma"  
/lab\_host="DH108 (phage-resistant)"  
/notes="Organ: skin; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."  
BASE COUNT 285 a 187 c 194 g 282 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.98e-91 Length: 948  
Score: 856.50 Matches: 186  
Percent Similarity: 89.67% Conservative: 5  
Best Local Similarity: 87.32% Mismatches: 14  
Query Match: 33.04% Indels: 8  
DB: 13 Gaps: 1

US-09-873-409-8 (1-514) x BM471690 (1-948)

QY 266 ValPheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285  
|  
|  
|  
Db 205 GTTTTCTTTAGTGTAAATCCATAGCAGTTATTGCAATGGAGCAGTCCTCCTCACTTTGAA 264  
|  
|  
|  
QY 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305  
|  
|  
|  
Db 265 ACCTTCGCAATAGCCGAGAGCTGCTTTTCATATTTTCCAGGTATTGATAGAAACCC 324  
|  
|  
|  
QY 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325  
|  
|  
|  
Db 325 AGTATAGATAAATTTTCCACAGCTGGATATAAACCTGAATCCATAGAGAACTGTGGAA 384  
|  
|  
|  
QY 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345  
|  
|  
|  
Db 385 TTTAAAAATGTTCTTCAATATTCATCAAGACCATCTATCAAGATTTCTGAAGGTCGT 444  
|  
|  
|  
QY 346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365  
|  
|  
|  
Db 445 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCTTGGTCGCTCTCAATGGCAGTGGGAAG 504  
|  
|  
|  
QY 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385  
|  
|  
|  
Db 505 AGTACGGTAGTCAGCTTCTGACAGAGTTATATGATCCGATGATGGCTTTATCATGGTG 564  
|  
|  
|  
QY 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVal 405  
|  
|  
|  
Db 565 GATGAGAAATGACATCAGAGCTTTAAATGTGCGGCAATATCGAGACCATATTTGGAGTGGT 624  
|  
|  
|  
QY 405 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 424  
|  
|  
|  
Db 625 TAGTCAAGAGCCCTGGTGTGTTTGGTTCGGGACCACCATCAGTAACAATATCAAGTATGGGACG 684



```

Qy 424 gAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 443
Db 685 AGATGATGGGACTGATGAACAATATGAAAAGGCAACAGGGAACCAAGGCGAATGAA 744
Qy 444 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaG1 462
Db 745 TTTTATCATGGAATTTCCCAAAAAAATTAATACATTGAAAGGGGAAAAAAGGACCTTCA 804
Qy 462 nMetSerGlyGlyGlnLysGlnArg 470
Db 805 AGGAATGGAAGGCCAAACCAACCAAG 829

RESULT 7
LOCUS BF584668 894 bp mRNA linear EST 12-DEC-2000
DEFINITION 602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
mRNA sequence.
ACCESSION BF584668
VERSION BF584668.1 GI:11658386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9798 row: 1 column: 02
High quality sequence stop: 651.

FEATURES
Location/Qualifiers
1..894
/organism="Mus musculus"
/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:4218385"
/lab_hosts="NCI_CGAP_Co24"
/notes="Organ: DHI0B (TI phage-resistant)"
/site 1: NotI;
/site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 263 a 203 c 242 g 185 t
ORIGIN
1 others

Alignment Scores:
Pred. No.: 1,85e-78 Length: 894
Score: 750.00 Matches: 156
Percent Similarity: 84.94% Conservative: 47
Best Local Similarity: 65.27% Mismatches: 36
Query Match: 28.94% Indels: 4
DB: 12 Gaps: 0

US-09-873-409-8 (1-514) x BF584668 (1-894)

Qy 266 ValPheSerValIleHisSerTyrCysTyrGlyAlaAlaValProHisPheGlu 285
Db 5 GTCTCTTTTCGGTGTAAAT-GGAGCATTCAGTGTGGACAGGCATCTCCAAATATTGAA 63
Qy 286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db 64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATATGATAAAGCC 123

```

```

Qy 306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db 124 AGTATAGACAGCTTCTCAAAAGAGTGGGCAACAAACACAGACATACAAAGAAATCTGGAA 183
Qy 326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysGlyLeu 345
Db 184 TTTTAAAGATATTTCACCTTCAGTTACCCATCTCGAAAAAGAAAGTTCCAGATCTTTGAAGGCCCTC 243
Qy 346 AspLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db 244 AATCTGAGGTGAAGACGGACAGACGGTGGCCCTGGTGGCAACAGTGGCTGTGGAAA 303
Qy 366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db 304 AGCACAACCTGCTCAGCTGATGCAAGGCTCTACGACCCCTAGATGGCTGTCAGTATC 363
Qy 386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db 364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTCGAGGAGATCATTTGTGTGGTG 423
Qy 406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db 424 AGTCAGAACTGTGTCTGTTGCCACACGATCGCGAGACATTCGCTATGCGCCGAGAA 483
Qy 426 AspValThrAspGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db 484 GATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGAGGCCAATGCTATGACTTTCATC 543
Qy 446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db 544 ATGAAATCGCCCAACCAATTTGACACCTCGTGGTGAGAGAGGGCGCAGCTGAGTGGG 603
Qy 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db 604 GGAACAGAAACAGAGATCGC-ATTGCCGGGGCCCTGGTCCGCAATCCCAAGTGCCTTTC- 661
Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 504
Db 662 GTGACGAGGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGTTCAGGCCGCA 717

RESULT 8
AK020318 872 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
DEFINITION library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
(P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION AK020318.1 GI:12860872
VERSION AK020318.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
clone lib:RIKEN full-length enriched mouse cDNA library
clone:9230106F14.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

```

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Ogawa, Y., Izawa, M., Matsura, S., Watahiki, M., Yoneda, Y., Ishikawa, T., Tanaka, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balzarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Maehama, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weltz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
AUTHORS

21085660  
11217851  
5 (bases 1 to 872)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Balzarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, T., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTITTTTITTTT 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTAAATAAATAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES  
source

Location/Qualifiers  
1..872  
/organism="Mus musculus"  
/db\_xref="FANTOM DB:9230106F14"  
/db\_xref="MGI:1911565"  
/db\_xref="taxon:10090"  
/clone="9230106F14"  
/sex="male"  
/tissue type="epididymis"  
/clone.lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
86..>872  
/note="data source:SPTR, source key:P21447, evidence:ISS putative similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3) (MDR1A)"  
/codon\_start=1  
/protein\_id="BAB32070.1"  
/db\_xref="GI:12860873"  
/db\_xref="MGI:1924956"  
/translation="MANSERTNGLOETNORYGPLEQVVKVGNQAVGPTEIPRFADNL DIVLMTGLIASMINGATVPLMSLVLGISIDHLINGCLVOTNRTKYQNCSONQEXKNE DIIVLTYYIGIGAAALIFGVQVSWFVITAAQTTRKQFFHSILAQDSWFGSD ICELNRTMTGDKLDCDGIPLMFQNSIGLVISLISLKWLSLVLSTSPLI MASSALCSRMIIISLTSKELDAYSKAGAAEEALSISIQTVTAFGA"  
BASE COUNT 262 a 178 c 187 g 245 t  
ORIGIN

Alignment Scores:

Pred. No.: 8,068-78 Length: 872  
Score: 744.50 Matches: 151  
Percent Similarity: 80.37% Conservative: 21  
Best Local Similarity: 70.56% Mismatches: 27  
Query Match: 28.72% Indels: 15  
DB: 11 Gaps: 2  
US-09-873-409-8 (1-514) x AK020318 (1-872)  
QY 1 MetIleLeuGlyTleLeuAlaSerLeuValAsnGlyValaCysLeuProLeuMetProLeu 20  
Db 230 ATGACTCTGGGAAATATTAGCATCCATGATAATATGGAGCCACCGTCTCTTAATGTCCCTG 289  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 290 GTTTAGGAGAAATAGTGCATCTTTAATTAATGATGCTCTAGTACAACTAACAGAACT 349  
QY 41 SerPhe-----PheArgLeuThr 46  
Db 350 AAATATCAGAACTGTTCTCAGACTCAAGAAAAGCTGAATGAGATATCATTTGTGTGACT 409  
QY 47 LeuTyrTyrValGlyTleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66  
Db 410 CTATATTATTATGGAATAGAGAGAGCTGCGCCCTCATTTTGGCTATGATGACAGATTTCTTC 459  
QY 67 TrpIleIleThrAlaAlaArgGlnThrTyrArgIleArgLysGlnPhePheHisSerVal 86  
Db 470 TGGGTCAATCTGCAGCCCGGCAACCAAGAACTCGAAACAGTTTTTTCATTCAATT 529  
QY 87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106  
Db 530 TTGGCAACAAGACATCAGCTGGTTTGTATGTCAGTGACATCTGTGAACCTTAACCCGCATG 589  
QY 107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125  
Db 590 ACTGGTGACATCAACAACTCTGTGATGTTATGGAGATAAGATCCCTCTGATGTTTCAG 649  
QY 126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145  
Db 650 AACATATCTGGGTTTCTATTGCGCTGGTGATGAAGTTTGATAAAGCTGGAACTCTCC 709  
QY 146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMet 165

Db	710	CTGGTGTTCTGTCTACATCTCTCTCATATGCTTTCATCGGACGTGCTCTAGCATG	769
Qy	166	ValleSerLeuThrSerLySGLuSerAlaTySerLySAlaGlyAlaValAlaGlu	185
Db	770	ATTATCTCATGTACCAACAGAGAGCTGGATGCTTATTCACAGCTGGGCTGTGGCTGAA	829
Qy	186	GlutValLeuSerSerIleArgThrValIleAlaPheArgala	199
Db	830	GAAGCCTTGTCATCAATCAACAGTCACAGCCTTTGGAGCC	871
RESULT	9		
LOCUS	AK014319		
DEFINITION	Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:3222401P09/ATP-binding cassette, sub-family B (MDR/TAP), member 8, full insert sequence.		
ACCESSION	AK014319		
VERSION	AK014319.1	GI:12852089	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-Format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staab, J. F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, N., Hume, D. A., Kariya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PUBMED	11217851		
REFERENCE	5 (bases 1 to 2676)		
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoaka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numata, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)		
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.		
FEATURES	Location/Qualifiers		
source	1..2676		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM DB:3222401P09"		
	/db_xref="MGI:1907488"		
	/db_xref="taxon:10090"		
	/clone="3222401P09"		
	/tissue type="head"		
	/dev_stage="14, 17 days embryo"		
gene	1..2676		
	/gene="Abcb8"		
CDS	72..2225		
	/gene="Abcb8"		
	/note="ATP-binding cassette, sub-family B (MDR/TAP), member 8"		
	data source:MGI, source key:MGI:1351667, evidence:ISS putative"		
	/codon_start=1		
	/protein_id="BAB29270.1"		
	/db_xref="GI:12852090"		
	/db_xref="MGI:1351667"		
	/translation="MLVHLFRFGIRGGVPVGSLSQSLRPFQTSAAKSSDRSLSSHLRLTYAQRVQLRAHLPRAPFASWSPACWVGTLVPAVLVWHPRLCLALCEAKESPQAQTRAPRLRFNFKLFWHLPHLLALGAALVQVPLPLLGQVPIVAKVYRDHMGFSVSSKLSVOLLVYVQGLLTFLVLSHIGERWMDNRKALFSSLLRODIAFFDAKTKTGLVSRLLTDDVQBFKSSFLVISOGLRSCVTQVLSVLSLRLTLMVAVPALMGVGTLMGSLGRKLSRQCEQIARATGVADALGNVTVRAFAMEKREERYQAELESCCCAEELGQVLFQGLSNIAFNQVLTGLTIFGSLVAGQQLKGGDLMSFLVASQTVQRSMASLVLFQVVRGLSAGARVFYFNALSPVPLTGTGCTIPNKDI		

RGSTFQNTSPYPCPGPNVLKFTLKLPSGKIVLVQSGGGKTTVASLLERYDP  
EAGSYTLGDHDLRTLPNQLQVIFISQBEVLFPATTIMENIRFGKLDASDEEYVTA  
AREANAHFISFPGDYSTVGERGTTLSGGQKQRLAIARALIKOFTVLILDEATSA  
DAESRVVQELDRASAGRTVLVIAHRSITVRAHSIIVMANGQVCEAGTHEELLKKG  
GLYSILRRQTLDAISLTSTPPAEKPEDPKSCQSKA"  
2653..2658  
/gene="Abcb8"  
/note="putative"  
2676  
/gene="Abcb8"  
/note="putative"

polyA\_signal

polyA\_site

BASE COUNT 544 a 785 c 756 g 591 t  
ORIGIN

Alignment Scores:

Pred. No.: 6.67e-73 Length: 2676  
Score: 710.00 Matches: 194  
Percent Similarity: 54.48% Conservativity: 98  
Best Local Similarity: 36.19% Mismatches: 184  
Query Match: 27.39% Indels: 60  
DB: 11 Gaps: 19

US-09-873-409-8 (1-514) x AK014319 (1-2676)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 471 ATTGTCTGGCTTAGTGGCGGCACTAGTGAAT-----GTGCAGATCCCTCTTG 518  
QY 21 ValLeuGlyGluMetSerAspAsnLeuIleSer-----GlyCysLeu 34  
Db 519 CTCCTGGGCGAGCTGGTAGATTGTCGCCAAGTACACGAGGAGCACATGGGGAGTTTC 578  
QY 35 ValGlnThrAsnThrTyrSerPheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54  
Db 579 GTGCTGAGTCCCGTAAGCTCAGCGTCCAGCTCTCTACTGTAC-----GGTGT 629  
QY 55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTyrIleThrAlaAlaArg 73  
Db 630 CAGGGAGCTGCTGACCTTTGGATACCTAGTCTGCTGTGCCACATT-----GGTGAGCGC 683  
QY 74 GlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyr 93  
Db 684 ATGGCCATGCATCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGAAGACATTTGCTTC 743  
QY 94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111  
Db 744 TTTGATGCCMAAAGACAGGGGAGCTAGTGCAGTCTGTACTACTGTATGTGCAGAAATTC 803  
QY 112 -----IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124  
Db 804 AAGTCATCTTCAAGCTTGTCATCTCCAGGAGCTG-----839  
QY 125 GlnAsnMetSerThrPheSerIleGlyValAlaValGlyLeu---ValLysGlyTyrLys 143  
Db 840 ---CGCAGCTGCACCCAGGTGATGTGGTGTAGCTGTGCTGTCTATGTCTCCCTGCG 896  
QY 144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaCysSer 163  
Db 897 CTTACCTGATGTGGCTGCTGCACACCCGCTCATGGGAGTGGGCACCTGTATGGGC 956  
QY 164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183  
Db 957 TCAGGCTCCGAAAGCTGTCTGCCAGTGTGCAGGAGCAGATTTGCCGGGCAACAGGTGTA 1016  
QY 184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203  
Db 1017 GCAGATGAGGCCCTTGGCAATGTTCGAGCTGTGCGGCCCTTCGCGATGGAGAGAGGAG 1076  
QY 204 LeuGlnArgTyrThrGlnAsnLeuLysAsp-----AlaLysAspPheGlyIle 219  
Db 1077 GAGGAGCGCTATCAAGCAGAACCTGGATGCTAGTGTGTAAAGCAGAAAGCTGGGCG--- 1133  
QY 220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe-----PheMetAsn 237

Db 1134 ---AGGGGATCGCTTGTTCAGAGGCTCTCAACATCGCTTTCAACTGTATGGTCTTG 1190  
QY 238 GlyThrTyrGlyLeuAlaPheThrTyrThrSerLeuIleLeuAsnGlyGlu---Pro 256  
Db 1191 GGCACC-----CTGTTTCATGGGGCTCCCTGTGGCTGCACAGCAGCTGAAA 1238  
QY 257 GlyThrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerTyrCys 276  
Db 1239 GGGGAGACCTCATGTCTCTCTGTGGCTTCCACAGACAGTACAGAGGTCTATGGCCAGC 1298  
QY 277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly-----AlaAla 294  
Db 1299 CTCCTGTCTCTTGGTCAG-----GTGTACGTGGGTGAGTGGCGGA 1343  
QY 295 PheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGly 314  
Db 1344 GCCCGAGTCTTCGATACATGAGCCCTGAGCCCTGTCTATCCCA-----TTGACCGGGGCG 1397  
QY 315 Tyr-----LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 332  
Db 1398 TACTGCATCCCAACAAGGACATTCGTGTTCCTCACCTTCCAAAATGTCACTTCAGC 1457  
QY 333 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 352  
Db 1458 TACCCCTGCAGACCTGGCTTCAATGTCTCAAGGACTTCACCTTGAAGCTGCCCTCTGGC 1517  
QY 353 GlnThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 372  
Db 1518 AAGATTGTGGCTCTTGTGGCCAGTCTGGGGGAGGAAACACACAGTTCCTCTTCTGCTG 1577  
QY 373 GlnArgLeuTyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAla 392  
Db 1578 GAACGCTTCTATGACCTGAAGCTGGCTCGTGACCTTGGATGGGCATGACCTGCGAACT 1637  
QY 393 LeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGluProValLeu 411  
Db 1638 CTCACCCCTCTCTGGCTCGGGGCCAGGTCTATAGTGTTCATCAGCCAGGAGCAGTCTG 1697  
QY 412 PheGlyThrThrIleSerAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 431  
Db 1698 TTTGCACCAACCATCATGAGAGATATCCGATTTGGGAAGCTGCTGCTCCGATGAGAG 1757  
QY 432 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 451  
Db 1758 GTGTACACAGCTGCACGAGAGCCAAATGCCAGAGTTCATCAGCAGCTTCCCGGATGGC 1817  
QY 452 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 471  
Db 1818 TACAGCACTGTGGTGTGGGCGGCGCACAACTTGTCTGTGGTGGCCAGAGCAGCGCCTA 1877  
QY 472 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 491  
Db 1878 GCCATCGCAGTGCCTCTCATCAGAGCCCAACAGTCTGCTATCTCTGACGAGGCGCACCA 1937  
QY 492 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 507  
Db 1938 GCGCTAGATGCAGAAATCCGAGAGGGTGGTACAGGAGGCCCTTGAGCCGG 1985

RESULT 10

BG298756

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

803 bp mRNA linear EST 21-FEB-2001  
602396681F1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:4511157 5',  
mRNA sequence.

BG298756

BG298756.1 GI:13063728

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 803)

NIH-MGC http://mgc.nci.nih.gov/.

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: csapbs-f@mail.nih.gov  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 Plate: LLAM10394 row: b column: 22  
 High quality sequence stop: 698.

**FEATURES**

source

1. .803 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4511157"  
 /clone\_lib="NIH MGC 94"  
 /tissue\_type="retina"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."  
 240 a 162 c 205 g 196 t

BASE COUNT

ORIGIN

## Alignment Scores:

Pred. No.: 3.63e-69 Length: 803  
 Score: 671.50 Matches: 149  
 Percent Similarity: 76.15% Conservative: 49  
 Best Local Similarity: 57.31% Mismatches: 55  
 Query Match: 25.91% Indels: 10  
 DB: 12 Gaps: 2

US-09-873-409-8 (1-514) x BG298756 (1-803)

QY 168 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 187  
 DB 15 TCATTTACTGATGAAGAACTCCATGCTTATGCAAAAGCTGGAGCAGTGTCTGAAGAATC 74  
 QY 188 LeuSerSerLeuArgThrValLeuAlaPheArgAlaGlnGluLysGluLeuGlnArgTyr 207  
 DB 75 TTAGCAGCCATCAGAACTGTGATTCCTTTGGA-GGACAAAGAGGAAGAACTTGAAGGTAC 133  
 QY 208 ThrGlnAsnLeuLysAspAlaLysAspPheGlyLeuLysArgThrLeuAlaSerLysVal 227  
 DB 134 AATAACAACCTTGAAGAAGCTAAAGGCTGGGATAAAGAAGCTATCACGGCCAAATC 193  
 QY 228 SerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGly 247  
 DB 194 TCCATGGTGCAGCTTTCTCCTTATCTATGTCATCATATGCTCTGGCATTTCTGGTATGG 253  
 QY 248 ThrSerLeuLeuAsnGlyGluProGlyTyrThrLeuGlyThrValLeuAlaValPhe 267  
 DB 254 ACTTCCTGGTCATCTCAAGAA-----TACTCTAT-GGACAAAGTCTCACTGCTTC 306  
 QY 268 PheSerValLeuHisSerSerTyrCysAlaGlyAlaValProHisPheGluThrPhe 287  
 DB 307 TTTTCCTGTTAATTTGGAGCATTCAGTGTGGACAGGCATCTCCAAATATTGAAGCCTTC 366  
 QY 288 AlaLeuAlaArgGlyAlaAlaPheHisLeuPheGlnValLeuAspLysLysProSerile 307  
 DB 367 GCCAATCCAGAGGAGCAGCTTATGAGTCTTCAAAATAATTGATAATAGCCCAAGTATA 426  
 QY 308 AspAsnPheSerThrAlaGlyTyrLysProGluSerLeuGluGlyThrValGluPheLys 327  
 DB 427 GACAGCTTCTCAAGAGTGGCCAAACCAACGACAAATACAAAGAAATCTGGAATTTAAG 486  
 QY 328 AsnValSerPheAntTyrProSerArgProSerLeuLysLeuLysGlyLeuAsnLeu 347

DB 487 AATATTCACTTCAGTTACCATCTCGAAAGAAGTTTCAGATCTTGAAGGCCCTCAATCTG 546  
 QY 348 ArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThr 367  
 DB 547 AGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGAAAGACACA 606  
 QY 368 ValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGlu 387  
 DB 607 ACTGTCCAGCTGATGCAAGAGCTTACGACCCCTA-GATGGCATGTCAGTATCGACGGA 665  
 QY 388 AsnAspIleArgAlaLeu-AsnValArgHis-TyrArg-AspHisIleGlyValValSer 406  
 DB 666 CAGGACATCAGAACCATCAATGTCAGGTATCTGAGGGGATCAT---GGGGTGGTCACT 722  
 QY 407 Gln-GluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 423  
 DB 723 CAAGGAACCTGGCTTGTGTTTGGCAACCCAGATCGCGAGAAAATTCGTATATGGG 774

## RESULT 11

AL520322

LOCUS AL520322 943 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL520322 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5  
 prime, mRNA sequence.

ACCESSION

AL520322

VERSION

AL520322.1

GI:12783815

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 943)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Location/Qualifiers

1. .943

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DB006YC15"

/clone\_lib="LTI\_NFL004\_NBC2"

/sex="male"

/tissue\_type="neuroblastoma cells"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT

253 a 225 c 253 g 211 t

ORIGIN

Alignment Scores:

Pred. No.: 1.47e-67 Length: 943

Score: 659.00 Matches: 131

Percent Similarity: 72.36% Conservative: 47

Best Local Similarity: 53.25% Mismatches: 66

Query Match: 25.42% Indels: 2

DB: 9 Gaps: 1

US-09-873-409-8 (1-514) x AL520322 (1-943)

QY 264 LeuAlaValPhePheSerValLeuHisSerSerTyrCysIleGlyAlaAlaValProHis 283

DB 2 CTGTATTATTTCAGTCTGTGTCTTTGTCATCGCGTGGGCAAGTCAGTTTCATTT 61

QY 284 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 303  
 Db 62 GCTCTGACTATGCAAGCAAAATATCAGCAGCCACATCATCATGATCATTTGAAAAA 121  
 QY 304 LysProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThr 323  
 Db 122 ACCCTTTGATTGACAGCTACAGCAGGAGGCTTAATCCGAAAGCAATTTGGAAGGAAT 181  
 QY 324 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysLys 343  
 Db 182 GTCACATTTGGTGAAGTTGATTCACTATCCACCCGCGACATCCAGCTGCTTCAG 241  
 QY 344 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 363  
 Db 242 GGCAGTGCAGCTGAGGTGAAGAGGCGCAGACGCTGGCTCTGGTGGCAGCAGTGGCTGT 301  
 QY 364 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIle 383  
 Db 302 GGAAGAGCAGACATGCTGCTGAGGCGGTCTTACGACCCCTTGGCAGGGAAGTG 361  
 QY 384 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 403  
 Db 362 CTGCTTGATGGCAAGAAATAAAGCAGCTGAATGTTTCAGTGGCTCCGACACACCTGGGC 421  
 QY 404 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 423  
 Db 422 ATGCTGCTCCAGAGGCCATCTGTTTGATGCTGCAGCATTTGCTGAGAATTCGCTATGGA 481  
 QY 424 ArgAspAsp-----ValThrAspGluGluMetGluArgAlaAlaAlaArgGluAlaAsnAla 441  
 Db 482 GACACAGCGGGTGTGTGCACAGGAAGAGATCGTCAGGGCGCAGCAAGAGGCGCAACATA 541  
 QY 442 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 461  
 Db 542 CATGCTTCATCGAGTCAGTGCCTTAATAATATAGACATAAAGTAGAGACAAAGAACT 601  
 QY 462 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 481  
 Db 602 CAGCTCTGCTGGTGGCCAGAAACCAACGATTCGCCATGCTGCTGCCCTTGTAGACAGCCT 661  
 QY 482 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 501  
 Db 662 CATATTTTGGTTTGGTGAAGCCAGCTCAGCTCTCGATACAGAAAGTGAAAGGTTGTC 721  
 QY 502 GlnAlaAlaLeuGluLys 507  
 Db 722 CAAGAAGCCCTGGACAAA 739

RESULT 12  
 AA243820 405 bp mRNA linear EST 06-AUG-1997  
 LOCUS z67g06.r1 Soares NHMPU S1 Homo sapiens cDNA clone IMAGE:668506 5',  
 DEFINITION similar to SW:MDRI\_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;,  
 mRNA sequence.

ACCESSION AA243820  
 VERSION AA243820.1 GI:1874631  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 405)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Scheilberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,  
 T., Waterston, R. and Wilson, R.  
 WashU-Merck EST Project 1997  
 TITLE Unpublished (1997)  
 JOURNAL Contact: Wilson RK  
 COMMENT Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800

Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Insert Length: 1572 Std Error: 0.00  
 Seq primer: -28ml3 rev2 ET from Amersham.

# FEATURES

source  
 1. .405  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5562473"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:668506"  
 /clone\_lib="Soares NHMPUs S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac  
 (Pharmacia) with a modified polylinker; Site 1: Not I;  
 Site 2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NBHM, pregnant uterus  
 NBHPU, and fetal heart NBHH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."

BASE COUNT 115 a 81 c 85 g 124 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.44e-67 Length: 405  
 Score: 654.00 Matches: 128  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.22% Mismatches: 0  
 Query Match: 25.23% Indels: 0  
 DB: 9 Gaps: 0

US-09-873-409-8 (1-514) x AA243820 (1-405)

QY 264 LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis 283  
 Db 18 CTCTCTGTTTCTTAGTGTAAATCATACAGTATTGTCATTGGACGACAGTCCCTCAC 77  
 QY 284 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 303  
 Db 78 TTTGAACCTTCGCAATAGCCCGAGGAGCTGCTTTCATATTTCCAGGTATTGATAAG 137  
 QY 304 LysProSerIleAspAsnPheserThrAlaGlyTyrLysProGluSerIleGluGlyThr 323  
 Db 138 AAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAAC 197  
 QY 324 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLysLys 343  
 Db 198 GTGGAATTTAAAAATGTTTCTTCAATTTATCCATCAAGACCATCTATCAAGATTCGAAA 257  
 QY 344 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 363  
 Db 258 GGTCTGAATCTCAGAAATTAAGTCTGGAGAGACAGTCCCTTGGTCGCTCTCAATGSCAGT 317  
 QY 364 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspGlyPheIle 383  
 Db 318 GGAAGAGTACGCTAGCTGCTGAGGCTTCTGAGAGGTTATATGATCCGATGATGCTTTATC 377  
 QY 384 MetValAspGluAsnAspIleArgAla 392  
 Db 378 ATGGTGGTGAATGACATCAGACGCT 404

RESULT 13  
 BU075336  
 LOCUS BU075336  
 DEFINITION BJO75336 NIBB Mochii normalized Xenopus tailbud library Xenopus  
 laevis cDNA clone XL057d10 5', mRNA sequence.

550 bp mRNA linear EST 11-DEC-2001  
 BJO75336  
 BJO75336 NIBB Mochii normalized Xenopus tailbud library Xenopus  
 laevis cDNA clone XL057d10 5', mRNA sequence.

1



US-09-873-409-8 (1-514) x BG587938 (1-780)

```
QY 277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 296
Db 3 TWGGGCAATCTGCACCGCATGGCTGCTATTTACAAAGCTAGATGGCGCTGTAAG 62
QY 297 IlePheGlnValIleAspLysPheSerIleAspAsnPheSerThrAlaGlyTyrLys 316
Db 63 ATTTCCGGATAATTGATCACCAGCCTGGTATAGATAGAAACAGTCAATCTGGATTGGAA 122
QY 317 ProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 336
Db 123 TTAGACACAGTACTGAGCTGTGTGAACCTGAAATATGTGGACTTCTTATCATCAAGA 182
QY 337 ProSerIleLysLysLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 356
Db 183 CTGAAGTCTGATCTCAATGATTTCTCTGAGTGTCTCCGCGAAAGACCATAGCT 242
QY 357 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyr 376
Db 243 TTAGTTGGTAGCAGCGCTCTGGCAAGACGACTGTGTCTTCTTATGAGAGATTCTAT 302
QY 377 AspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 396
Db 303 GATCCAACTTCAGGCAAGTAATGTTGGATGGCGCATGACATTAATACTTGAAGCTTAA 362
QY 397 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIle 416
Db 363 TGGTTGAGGCAACAATAGACGACTGAGTGGCAAGACCTGCTTTGTTGCCACCATGAT 422
QY 417 SerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAla 436
Db 423 CGAGAAATATATCTCTGGGAGCGCTGATGCAACACGAGTTGAGATTGAAGAAGCTGCT 482
QY 437 ArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuVal 456
Db 483 AGGTTGCTAAATCTCATCTTATCATCATCAAACTTCTCGAAGGCTTTGAAACTCAGSTA 542
QY 457 GlyGlyLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAla 476
Db 543 GGAGAAAGAGGACTACAACTTTCTGGAGGCAAAAACAAGAATAGCAATAGCAAGGGCA 602
QY 477 LeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGlu 496
Db 603 ATGCTAAAAATCCAGCAATTTCTTCTCTGATGAGGCAACAGTGCATTTGACTCTGHA 662
QY 497 SerLysSerAlaValGlnAlaLeuGluLys 507
Db 663 TCAGAAAAAGCTGGTGAAGAAGCACTTGACCG 695
```

## RESULT 15

```
BF692596
LOCUS 602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
DEFINITION mRNA sequence.
```

ACCESSION BF692596

VERSION BF692596.1

KEYWORDS EST.

SOURCE human.

## ORGANISM

```
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

1 (bases 1 to 559)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: ATCC/DCTD/DTP

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM1203 row: a column: 13

High quality sequence stop: 555.

## FEATURES

Location/Qualifiers

1..559

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4334100"

/clone\_lib="NIH MGC 62"

/tissue\_type="melanotic melanoma, high MDR"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ; Vector: pDMR-LIB (Clontech); Site\_1: SfiI (ggccgctcgcc); Site\_2: SfiI (ggccattggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 161 a 118 c 141 g 139 t

## ORIGIN

Alignment Scores:

Pred. No.: 7,08e-64 Length: 559

Score: 625.00 Matches: 134

Percent Similarity: 96.43% Conservatve: 1

Best Local Similarity: 95.71% Mismatches: 4

Query Match: 24.11% Indels: 3

DB: 12 Gaps: 0

US-09-873-409-8 (1-514) x BF692596 (1-559)

```
QY 376 TyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db 1 TATGATCCGATGATGGCTTTATCAGGTGGATGAGATGACATCAGAGCTTTAAATGTG 60
QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db 61 CGCATTTATCGAGACCATAT-GGAGTGGTTAGTCAAGAGCCTGT-TTGTTCGGGACAC 118
QY 416 IleSerAsnAsnIleLysTyrGlyArgAspValThrAspGluGluMetGluArgAla 435
Db 119 ATCAGTAAACATATCAAGTATGACGAGATGATGTGACTGATGAAGAGATGGAGAGCA 178
QY 436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db 179 GCAAGGGAAGCAATGCATATGATTTTATCATGGAGTTTCTTAATAAATTAATCATTTG 238
QY 456 ValGlyGluLys-GlyAlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaAr 475
Db 239 GTAGGGGAAAACAGGAGCTCAATAGTGGAGGCGGAGAAACAGAGGATCGCAATTGCTCG 298
QY 475 gAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSe 495
Db 299 TGCCTTAGTTCGAAACCCCAAGATTTCTGATTTAGATGAGGCTACGCTGCTGCTGATTC 358
QY 495 rGluSerLysSerAlaValGlnAlaLeuGluLysAspThrProArgTyrSerPhe 514
Db 359 AGAAACGAGTACGCTGTTCAAGTCTGAGGAGAGGATACCCCGAGGTATTCATTT 416
```

Search completed: March 31, 2003, 13:51:49

Job time : 3547.51 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 02:24:20 ; Search time 75.2727 Seconds  
(without alignments)  
2094.145 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592

Sequence: 1 MILGILASLVNGACLPMLP.....SESKSAVOALEKDPTRYSP 514

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09873409/runat\_27032003\_115422\_19259/app\_query.fasta\_1.7544  
-DB=Issued Patents NA -Qfmt=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09873409 @CGN 1 1 578 @runat\_27032003\_115422\_19259 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCUTS\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1576	60.8	4233	3	US-09-120-513-1
2	1576	60.8	4233	3	US-09-450-105-1
3	1576	60.8	4669	2	US-08-752-447-1
4	1576	60.8	4669	4	US-09-316-167-1
5	1570	60.6	4646	1	US-08-181-471-2
6	1570	60.6	6505	2	US-08-793-610-5
7	1570	60.6	9318	2	US-08-793-610-6
8	1568.5	60.5	4264	2	US-08-784-649A-1
9	1568.5	60.5	4264	2	US-08-784-649A-5
10	1568	60.5	4669	6	5206352-3
11	1549	59.8	4669	2	US-08-583-276-18
12	975.5	37.6	4224	1	US-08-612-521-1

13	959	37.0	4047	2	US-08-612-734B-1	Sequence 1, Appli
14	939.5	36.2	4800	2	US-08-612-734B-3	Sequence 3, Appli
15	937.5	36.2	4002	2	US-08-996-545-1	Sequence 1, Appli
16	937.5	36.2	4002	2	US-08-996-545-3	Sequence 3, Appli
17	937.5	36.2	4002	4	US-09-328-320-1	Sequence 1, Appli
18	937.5	36.2	4002	4	US-09-328-320-3	Sequence 3, Appli
19	920	35.5	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	902.5	34.8	2726	1	US-08-461-823-1	Sequence 1, Appli
21	859	33.1	6143	1	US-08-612-521-3	Sequence 3, Appli
22	809	31.2	3792	1	US-09-351-224E-10	Sequence 10, Appli
23	770	29.7	3924	2	US-08-996-644-3	Sequence 3, Appli
24	770	29.7	3924	2	US-09-352-552-3	Sequence 3, Appli
25	770	29.7	3927	2	US-08-996-644-1	Sequence 1, Appli
26	770	29.7	3927	3	US-09-352-552-1	Sequence 1, Appli
27	712	27.5	3999	4	US-09-351-224E-9	Sequence 9, Appli
28	669.5	25.8	2376	1	US-08-394-880B-1	Sequence 1, Appli
29	641	24.7	2061	4	US-09-061-764A-17	Sequence 17, Appli
30	639.5	24.7	3909	1	US-08-232-537-1	Sequence 1, Appli
31	635	24.5	1959	4	US-09-061-764A-4	Sequence 4, Appli
32	598	23.1	2244	4	US-09-061-764A-18	Sequence 18, Appli
33	581	22.4	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
34	543.5	21.0	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
35	529	20.4	7760	4	US-08-961-527-63	Sequence 63, Appli
36	521	20.1	5120	3	US-08-772-270A-6	Sequence 6, Appli
37	521	20.1	8370	2	US-08-488-706-1	Sequence 1, Appli
38	498.5	19.2	13188	4	US-08-961-527-70	Sequence 70, Appli
39	498.5	19.2	441529	4	US-09-103-840A-1	Sequence 1, Appli
40	497.5	19.2	7721	3	US-08-772-270A-14	Sequence 14, Appli
41	496.5	19.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
C 42	495.5	19.1	28804	2	US-08-592-874-1	Sequence 1, Appli
C 43	495.5	19.1	28804	3	US-09-096-942-2	Sequence 2, Appli
C 44	495.5	19.1	28804	3	US-09-096-867-2	Sequence 2, Appli
C 45	472.5	18.2	2407	2	US-08-895-522-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-120-513-1  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Cheney, Richard  
; APPLICANT: Ellens, Harma  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
; TITLE OF INVENTION: SCREENING METHODS THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY:  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,513  
; FILING DATE: 22-JUL-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GPO0008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4233 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

US-09-120-513-1

Alignment Scores:

Pred. No.: 2,88e-179 Length: 4233

Score: 1576.00 Matches: 308

Percent Similarity: 75.90% Conservative: 92

Best Local Similarity: 58.44% Mismatches: 105

Query Match: 60.80% Indels: 22

DB: 3 Gaps: 4

US-09-873-409-8 (1-514) x US-09-120-513-1 (1-4233)

Qy 1 MetileLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db 170 ATGGCTCTGGGAATCTCGCTGCTATCATCCACGGAACCCCTGCTCCCTCCTGATGCTG 229

Qy 21 ValLeuGlyGluMetSerAspAsnLeuIle----- 30

Db 230 GTGTTCGGATACATGACAGATAGTTTACCAAGCAGAGACCCCATCTGCCGAGGCTT 289

Qy 31 -----SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu----- 45

Db 290 ACTAATCAAGTGAATCAACAGTACACAGACCGCTCAGCGACAGCAGCTCTGGAGGAGAC 349

Qy 46 -----ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db 350 ATGGCCATGTACGCTACTATTACCGGCAATGGTCCCGGTGCTCATCTGTTGCCCTAC 409

Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgGln 81

Db 410 ATCCAGGTTTCACTTGGTCTGGCAGCTGGGAGACAAATACACAGATTTAGCAGAAG 469

Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101

Db 470 TTTTTCATGCTCATCAATCAGAGATAGCTGGTTGACGTGAATGACGCTGGGGAG 529

Qy 102 LeuAsnThrArgMetThr---AspIleAspIlyIleSerAspGlyIleGlyAspIlyIle 120

Db 530 CTCACACCCCGGCTCAGATGACGTCTCCAAATTAATGACGGAATGGTGACAAACTT 589

Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140

Db 590 GGAATGTTCTTTCAGTCCATPACACACATTTTACCCCGGTTTATAATAGGATTTAAGT 649

Qy 141 GlyTrpLysLeuThrLeuValThrSerThrSerProLeuIleMetAlaSerAlaAla 160

Db 650 GGTGTGAAGCTAACCCTTGTAATTTGGCCGTCAGCCCTCTTATTGGGTGCTCATCTGCC 709

Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180

Db 710 ATGTGGCCAAAGGACTGACTTCTTACTAATAAGGAACCTCCAGGCTTATGCGAAAGCT 769

Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200

Db 770 GGACAGTGTCCGAGGAGGCTTTAGCAGCCATCAGACTGTGATGGTTGGAGGACAA 829

Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyLys 220

Db 830 AAGAAGGAACCTTGAAGGTACAATAAAATTTAGAAGAGCTAAAGAGTTGGCATAAAG 889

Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyTyr 240

Db 890 AAGCCATCAGCGCAACATTTCCATAGGTATTCCTACCTGTTGGTCTATGCGTCTTAT 949

Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260

Db 950 GCATG3CATTTCTGGTATGGACCTCTTGCTCTCAATGAA-----TATTCATT 1003

Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280

Db 1004 GGACAGTGTCTTACCGTCTTCTTCTATTTATTGGGACTTTTCAGTATGGACATTTA 1063

Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300

Db 1064 GCCCAAAACATAGAAGCCTTTGCAATGCAAGAGGGGACGCCCTATGAAATCTTCAAGATA 1123

Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320

Db 1124 ATTGATAATGAGCCAGCAATTTGACAGCTTCTCAACCAAGGACACAAACACAGACATATA 1183

Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340

Db 1184 ATGGAAATTTGGAATTTAAATAATGTTTACTTCAACTACCCATCAGCAAGTGAAGTTAAG 1243

Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360

Db 1244 ATCTGAAGGCCCTCAACCTGAAGGTGAAGCGGCGAGACCGTAGCCCTGGTTGGCAAC 1303

Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyrAspProAspAsp 380

Db 1304 AGTGGCTGTGGGAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363

Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400

Db 1364 GGCAGGTGCTAGTATCAGCGACAGACATCAGGACCATCAATGTGAGGTATCTGCGGAA 1423

Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420

Db 1424 ATCATTTGGGTGTGTAGTCAGGAACCCGCTGTTGCCACCGATTCGCGAAACATT 1483

Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440

Db 1484 CGCTATGGCCGAGAAACAGTCACCATGGATGATAGAGAGAAGCTGTCAAGGAAGCCAT 1543

Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460

Db 1544 GCCTATGACTTCATCATGAACTGCCCCCAAAATTTAAACACCCCTGGTTGGTGAGAGAGG 1603

Qy 461 AlaGluMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480

Db 1604 GCGCAGTGTAGTGGGACAGAAACAGAGATCGCCATTCGCCGGCCCTGGTCCGCAAC 1663

Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500

Db 1664 CCCAAGATCCTTTTGTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGCGTG 1723

Qy 501 ValGlnAlaAlaLeuGluLys 507

Db 1724 GTTCAGCGCGCTCTGGATAG 1744

#### RESULT 2

US-09-450-105-1

; Sequence 1, Application US/09450105

; Patent No. 6169166

; GENERAL INFORMATION:

; APPLICANT: Kimberly Anne Brun

; APPLICANT: Richard James Chenery

; APPLICANT: Harma Ellens

; APPLICANT: John Anthony Feild

; APPLICANT: Lin Yue

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES

; FILE REFERENCE: GP-50008-DI

; CURRENT APPLICATION NUMBER: US/09/450,105

; EARLIER APPLICATION NUMBER: 09/120,513

; EARLIER FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 4233  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
US-09-450-105-1

## Alignment Scores:

Pred. No.: 2, 88e-179 Length: 4233  
Score: 1576.00 Matches: 308  
Percent Similarity: 75.90% Conservative: 92  
Best Local Similarity: 58.44% Mismatches: 105  
Query Match: 60.80% Indels: 22  
DB: 4 Gaps: 4

US-09-873-409-8 (1-514) x US-09-450-105-1 (1-4233)

```
QY 1 MetIleuGlyIleuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 170 ATGGCTCTGGGAACCTCGCTGCTATCATCCACGGAACCTCGCTCCCTCTCGTATGCTG 229
QY 21 ValIleuGlyCysLeuMetSerAspAsnLeuLeu----- 30
DB 230 GTCTTCGGATACATGACAGATAGTTTATCCCAAGCAGAGACCCGCATTCGCGAGCGTT 289
QY 31 -----SerGlyCysLeuValGlnThrAsnThrTy-SerPheArgLeu----- 45
DB 290 ACTAATCAAGTGAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349
QY 46 -----ThrLeuTyTrpValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 350 ATGGCCATGTACGCCCTACTATTACAGGGCATTTGGTGCCTGCTCATCGTTGCTTAC 409
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrIleArgIleGlyGln 81
DB 410 ATCCAGGTTTCACTTTGGTGGTGGGAGTGGGAGCAATAACACAGATTAAGCGCAAG 469
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 470 TTTTTCATGCCATCATGATCAGGAGATAGGCTGGTTGACGTGATGACGCTGGGGAG 529
QY 102 LeuAsnThrArgMetThr---AspIleAspIleSerAspGlyIleGlyAspIle 120
DB 530 CTCAACACCCGGCTCAGATGACGCTCCAAATAATGACGGAATGGTGACAAACTT 589
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 590 GGAATGTCTTTTCAGTCCATAACGACATTTTCAGCCGGTTTATAATAGGATTTATAAGT 649
QY 141 GlyTrpIleLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 650 GGTGGAAGCTAACCTTGAATTTTGGCGTCAGCCCTCTTATTTGGTGTTCATCTGCC 709
QY 161 AlaCysSerArgMetValIleLeuThrSerLysGluLeuSerAlaTySerLysAla 180
DB 710 ATCGGCGAAGGCTACTGACTCTATTACTAATAAGGAACCTCCAGGCTTATGCGAAGCT 769
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 770 GGAGCAGATTGCCGAAGAAGCTTTAGCAGCCATCAGAACTGTGATGGTTTGGAGGACAA 829
QY 201 GluIleGluLeuGlnArgTyThrGlnAsnLeuIleAspAlaIleAspPheGlyIleLys 220
DB 830 AAGAAGGAACCTTGAAGGTACAATAAATAATTAGAAGAGCTAAAGAGTTGGCATAAG 889
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyPhePheMetAsnGlyThrTy 240
DB 890 AAGCCATCAGGCCCAACATTTCCATAGGTATTGCCCTACCTGTGCTATGCTCTTAT 949
QY 241 GlyLeuAlaPheTrpTyTrpGlyThrSerLeuIleAsnGlyGluProGlyTyThrIle 260
DB 950 GCACCTGGCATCTTGGTATGGGACCTCCTTGGTCTCTCAAAATGAA-----TATTCTATT 1003
```

```
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyTrpCysIleGlyAlaAla 280
DB 1004 GGACAAGTCTTACCGCTCTTCTCTCTATTTATTGGGACTTTTCAGTATTGGACATTTA 1063
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
DB 1064 GCCCCAAACATAGAAGCTTTGCAATGCAAGAGGGGAGCCCTATGAATCTTCAAGATA 1123
QY 301 IleAspLysProSerIleAspAsnPheSerThrAlaGlyTyTrpLysProGluSerIle 320
DB 1124 ATGTATAATGAGCCCAAGCATTTGACAGCTTCTCAACCAAGGACACAAACACAGACATTA 1183
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyTrpSerArgProSerIleLys 340
DB 1184 ATGGGAATTTGGAAATTTAAAAATGTTTCACTTCAATACCCCATCACGAAATGAAGTTAAG 1243
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1244 ATCTTGAAGGGCTCAACCTGAAGGTGAAGAGCGGACAGCGTAGCCCTGGTTGGCAAC 1303
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyTrpAspProAspAsp 380
DB 1304 AGTGGCTGTGGGAAAGACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp 400
DB 1364 GCGAGGTCTAGTATCGACGACGAGGACATCAGGACCATCAATGTGAGGTATCTGCGGAA 1423
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1424 ATCATTTGGGGTGGTGTAGTCAGGAACCCGCTGCTTTGCCACCAAGATTCGCGAAACAT 1483
QY 421 LysTyTrpGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1484 CGCTATGGCGGAAACCTCACCATGATGATAGAGAAAGCTCTCAGGAGGCAAT 1543
QY 441 AlaTyTrpPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
DB 1544 GCCTATGACTTTCATCATGAACTGCCCCCAAAATTTAACACCCCTGTTGTGTGAGAGGG 1603
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 1604 GCGACCTAGTGGGGACAGAAACAGAGGATCGCATTTGCCCGGCCCTGGTCCGCAAC 1663
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 1664 CCAAGATCTTTTGTGTGATGAGGCCCACTTGGACACAGAAAGCGGCGTG 1723
QY 501 ValGlnAlaAlaLeuGlyLys 507
DB 1724 GTTCAGCGGCTCTGGATAAG 1744
```

## RESULT 3

US-08-752-447-1

; Sequence 1, Application US/08752447

; Patent No. 5994088

; GENERAL INFORMATION:

; APPLICANT: Mechtner, Eugene

; APPLICANT: Roninson, Igor B

; TITLE OF INVENTION: Methods and Reagents for Preparing and

; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert &amp; Berghoff Ltd.

; STREET: 300 South Wacker Drive, Seventh Floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/752,447

FILING DATE: 15-NOV-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 599408nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 95,1121

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-9808

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4669 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: 5'UTR

LOCATION: 1..424

FEATURE:

NAME/KEY: CDS

LOCATION: 425..4264

FEATURE:

NAME/KEY: 3'UTR

LOCATION: 4265..4669

US-08-752-447-1

#### Alignment Scores:

Pred. No.: 3 4e-179 Length: 4669  
Score: 1576.00 Matches: 305  
Percent Similarity: 77.04% Conservative: 101  
Best Local Similarity: 57.87% Mismatches: 99  
Query Match: 60.80% Indels: 22  
DB: 2 Gaps: 5

US-09-873-409-8 (1-514) x US-08-752-447-1 (1-4669)

Qy 1 MetLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 575 ATGGTGGGGAACCTTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 634  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 635 GTGTTGGAGAAATGCACAGATATCTTTGCAAAATCGAGAAATTTAGAAAGATCTGATGCA 694  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 695 AACATCACTAATAGAGTGATATCAATGATACAGGTTCTTCATGAATCTGGAGAAAC 754  
Qy 44 -----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 755 ATGACAGGATATGCTATTAATACAGTGAATTTGGTGGGGTGGTGGTGGCTTAC 814  
Qy 62 IleGlnIleSerLeuTyrPheIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81  
Db 815 ATTCAGGTTTCATTTGGTGGCTGGCAGCTGGAAGACAAATACACAAATTTAGAAACAG 874  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101  
Db 875 TTTTTCATGCTATATCCGACAGAGATAGCTGGTTGATGTCACGATGTTGGGAG 934  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 935 CTTAACACCCGATTCACAGATGATCTCTCCAGATTAATGAAGAAATTTGGTGACAAAT 994  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 1054

Qy 141 GlyTyrPlysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 1055 GGTGTGAAGCTAACCTTTGTGATTTTGGCCATCAGCTCTGTTCTTGACGTGTCAGCTGCT 1114  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 1115 GTCCTGGCAAGATACCTTCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 1175 GGAGCAGTAGCTGAAGAGCTCTTGGCAGCANTTAGAACTGTGATTCATTTGGAGGACAA 1234  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 1235 AAGAAAGAACTTGAAGGTACAAACAAATTTAGAAAGAGCTAAAGAAATTTGGGATAAG 1294  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCTTAT 1354  
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 260  
Db 1355 GCTGTGGCTTCTGTGATGGACCACTTGTCTCTCAGGGAA-----TATCTATT 1408  
Qy 261 GlyThrValLeuAlaValPheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 1409 GGACAGTACTCACTGTATTTCTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1468  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1469 TCTCCAAGCATTTGAAGCATTTGCAAAATGCAAGAGAGCAGCTTATGAAATCTTCAAGATA 1528  
Qy 301 IleAspLysIysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1529 ATTGATAAATGAAGCAAGTATTGACAGCTATTCGAAGTGGGCACAAACAGATAATATT 1588  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 1589 AAGGAAATTTGGAATTCAGAAATGTTCACTTCACTTACCCATCTCGAAAGAAAGTTAAG 1648  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1649 ATCTGAAGGCCCTGAACCTGAAGTGCAGAGTGGCAGACGGTGGCCCTGGTGGAAAC 1708  
Qy 361 AsnGlySerGlyLysSerThrValGlnLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1709 AGTGCTGTGGGAAGACACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1768  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1769 GGGATGTCAGTGTTCATGGACAGGATATTAGGACCATAAATGTAAAGTTTCTACGGAA 1828  
Qy 401 HisIleGlyValValSerGlnUpProValLeuPheGlyThrThrIleSerAsnIle 420  
Db 1829 ATCATTTGGTGTGGTGCAGGAACCTGTATTGTTGTCACCCACGATAGCTGAAACATT 1888  
Qy 421 LysTyrGlyArgAspValThrAspGluLeuMetClnuArgAlaAlaArgGluAlaAsn 440  
Db 1889 CGCTATGGCGTGAATGTCCATGATGAGATGAGAAAGCTGTCAAGAGAGGCCAAT 1948  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyLysGly 460  
Db 1949 GCCTATGATTTTATCATGAACCTGCCTCAATAATTGACACCTGTTGGTGGAGAGAGGG 2008  
Qy 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 2009 GCCCAGTTGATGGTGGCAGAAAGCAGAGGATGCCATTGCGTGGCTTGGTTCCGAAC 2068  
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 2069 CCCAAGATCTCTCTGCTGGATGAGGCCACCTGACGCTTGGACACAGAAAGCAGACGAGT 2128  
Qy 501 ValGlnAlaAlaLeuGluLys 507

Db	2129	GTTCAGGTGGCTCTGGATAAG	2149
<p>RESULT 4</p> <p>US-09-316-167-1</p> <p>Sequence 1, Application US/09316167</p> <p>Patent No. 6365357</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Mechethner, Eugene</p> <p>APPLICANT: Roninson, Igor B</p> <p>TITLE OF INVENTION: Methods and Reagents for Preparing and</p> <p>TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein</p> <p>NUMBER OF SEQUENCES: 2</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: McDonnell Boenhen Hulbert &amp; Berghoff Ltd.</p> <p>STREET: 300 South Wacker Drive, Seventh Floor</p> <p>CITY: Chicago</p> <p>STATE: Illinois</p> <p>COUNTRY: USA</p> <p>ZIP: 60606</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent in Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/09/316.167</p> <p>FILING DATE:</p> <p>CLASSIFICATION:</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 08/752,447</p> <p>FILING DATE: 15-NOV-1996</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: No.'6365357nan, Kevin E</p> <p>REGISTRATION NUMBER: 35,303</p> <p>REFERENCE/DOCKET NUMBER: 95,1121</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 312-913-0001</p> <p>TELEFAX: 312-913-9808</p> <p>INFORMATION FOR SEQ ID NO: 1:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 4669 base pairs</p> <p>TYPE: nucleic acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: cDNA</p> <p>FEATURE:</p> <p>NAME/KEY: 5'UTR</p> <p>LOCATION: 1..424</p> <p>FEATURE:</p> <p>NAME/KEY: CDS</p> <p>LOCATION: 425..4264</p> <p>FEATURE:</p> <p>NAME/KEY: 3'UTR</p> <p>LOCATION: 4265..4669</p> <p>US-09-316-167-1</p> <p>Alignment Scores:</p> <p>Pred. No.: 3,4e-179 Length: 4669</p> <p>Score: 1576.00 Matches: 305</p> <p>Percent Similarity: 77.04% Conservative: 101</p> <p>Best Local Similarity: 57.87% Mismatches: 99</p> <p>Query Match: 60.80% Indels: 22</p> <p>DB: 4 Gaps: 5</p> <p>US-09-873-409-8 (1-514) x US-09-316-167-1 (1-4669)</p> <p>Qy 1 MettleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20</p> <p>Db 575 ATGTGGTGGGAACCTTTGGCTGCCATCATCGATGGGCTGGACTTCTCTCATGATGCTG 634</p> <p>Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31</p>			



```
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGGATGTCAGTGTGTAGGACAGGATATTAGGACCAATAAATGTAAGGTTTCTACGGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1829 ATCATTTGGTGTGGTGCAGTCAGGAACCTGTATTGTTGCCACCACGATAGCTGAAACATT 1888
QY 421 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1889 GCGTATGGCGGTGAAATGTCACCATGGATGAGATTGAGAAAGCTGCAAGGAAGCCCAAT 1948
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1949 GCCTATGACTTATCATGAACCTGCTCATTAATTTGACACCCCTGGTTGGAGAGAGGG 2008
QY 461 AlaGlnMetSerGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2009 GCCCAGTTGAGTGTGGCAGAGCAGAGGATGCCATTGCGCTGCCCTGGTTGCGCAAC 2068
QY 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAGCAGTG 2128
QY 501 ValGlnAlaAlaLeuGluLys 507
Db 2129 GTTCAGGTGGCTCTGGATAG 2149
RESULT 5
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
```

```
ANTI-SENSE: NO
FEATURE: CDS
NAME/KEY: CDS
LOCATION: 425..4267
US-08-181-471-2
Alignment Scores:
Pred. No.: 1,79e-178 Length: 4646
Score: 1570.00 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 1 Gaps: 5
US-09-873-409-8 (1-514) x US-08-181-471-2 (1-4646)
QY 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 575 ATGGTGGTGGGAACCTTTGGCTGCCATCATCATCTGGGGCTGGAGCTCTCTCATGATGCTG 634
QY 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 635 GTGTTTGGGAATGACAGATATCTTTGCCAATGTCAGGAAATTTAGAAAGATCTGATGTCA 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 695 AACATCACTAATAGAAAGTGATATCAATGATATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 755 ATGACCAGGTATGCCATTATTATACAGTGAATTTGGTGGGGTGTGGTTGCTGCTTAC 814
QY 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 815 ATTCAGGTTTCATTTTGGTCCCTGGCAGCTGGAGACAAATACACAAATTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGTATGTCACGATGTTGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 935 CTTAACAACCCGACTTACAGATGATGCTCTTAAGATTAAATGAAGTTATTGGTGCAAAATT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 995 GGAATGTTCTTTTCAGTCAATGGCAACATTTTCACGGGTTTATAGTAGGATTACACGT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 1055 GGTTGGAAGCTAACCTTGTGATTTTGGCCATCAGTCTCTTCTGGACTGTGAGCTGCT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 1115 GTCTGGGCAAGATACCTTCTTACTGTATAAAGAACTCTTAGCGTATGCATAAGCT 1174
QY 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 1175 GGAGCAGTAGCTGAAGAGGCTCTGGCAGCAATTAGAATCTGTGATTCATTTGAGAGCAA 1234
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db 1235 AAGAAAGAACTTCGAAAGGTACAAACAAAAATTTAGAAAGAGCTTAAAGAAATGGGATAAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1295 AAAGCTATTACAGCCCAATATTTCTATAGGTGCTGCTTCTCTGCTGATCTATGATCTTAT 1354
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db 1355 GCTCTGGCCCTTCGGTATGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1408
QY 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
```

```
Db 1409 GGACAGTACTCACTGATCTCTTCTGTTAATTTGGGCTTTTAGTGTGGACAGGCA 1468
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyValAlaPheHisIlePheGlnVal 300
Db 1469 TCTCCAAGCATTTGCAATTTGCAATGCAAGGAGGAGCTTATGAATCTTCAAGATA 1528
Qy 301 IleAspLysPheProSerIleAspPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1529 ATTGATAAAGCCAGTATTGACAGTATTGACAGTATTGACAGTGGGCACAAACAGATATATT 1588
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1589 AAGGGAATTTGGAATTCAGAATGTTCTACTTCAGTATCCCATCTCGAAAGAGTTAAG 1648
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1649 ATCTTGAAGGGCTGAACCTGAAGGTGCAGATGGGCAGAGCGTGGCCCTGGTTGGAAC 1708
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1768
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1769 GGGATGCTCAGTGTTCATGGACAGGATATTAGGACCATATAATGTAAGGTTTCTACGGAA 1828
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db 1829 ATCAITGGTGTGGTGTGAGTCCAGGAACCTGTATTGTTCCACACCATAGTCTGAACATTT 1888
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1889 CGCTATGGCGGTGAAATGTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1948
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1949 GCCTATGACATTTATCATGAACCTGCCTCATATAATTTGACACCCCTGGTTGGAGAGAG 2008
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 2009 GCCAGTTGAGTGTGGGCAAGAGCAGAGATGCCATTCACGTCCCTGGTTGCCAAC 2068
Qy 481 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCAGCAGTG 2128
Qy 501 ValGlnAlaLeuGluLys 507
Db 2129 GTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 6

```
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; US-08-793-610-5

Alignment Scores:
Pred. No.: 3,17e-178 Length: 6505
Score: 1570.00 Matches: 305
Percent Similarity: 76.85% Conservative: 100
Best Local Similarity: 57.87% Mismatches: 100
Query Match: 60.57% Indels: 22
DB: 2 Gaps: 5

US-09-873-409-8 (1-514) x US-08-793-610-5 (1-6505)
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyValAlaCysLeuProLeuMetProLeu 20
Db 1967 ATGGTGTGGGAACTTGGCTGCCATCATCCATGGGGCTGGACTTCTCTCATGATGCTG 2026
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 2027 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAAATTTAGAGATCTGATGTCA 2086
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 2087 AACATCACTAATAGAGTATATCAATGATACAGGGGTTCTTCAATGAATCTGGAGGAAGAC 2146
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 2147 ATGACCAAGTATGCCATTATTACAGTGAATTTGGTGGGGTGTGGTGTGGTGTGGTGTAC 2206
Qy 62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 2207 ATTCAGTTTTCATTTTGGTGGCTGGCAGCTGGAGACAAATACACAAATTTAGAAAACAG 2266
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 2267 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGCTGTTGATGTCACCGATGTTGGGGAG 2326
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 2327 CTTAACACCCGACTACAGATGATGCTCTTAAGATTAAAGAGTTATTGTCACAAATTT 2386
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 2387 GGAATGTTCTTCAGTCAATGGCAACATTTTTCACGTGGGGTTTATAGTACAGATTACAC 2446
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
```

Db 2447 GGTTGGAAGCTAACCCCTGTGATCTTTGGCCATCAGTCTCTCTGACCTGTGACGCTGCT 2506  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerIleGluLeuSerAlaTySerIleAla 180  
Db 2507 GTCTGGCAAGATACCTTCTTCTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2566  
Qy 181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 2567 GGAGCACTAGCTGAAGAGGCTTGGCAGCAATAGAACTGTGATGCTTTGGAGGACAA 2626  
Qy 201 GluIysGluLeuGlnArgTyThrGlnAsnLeuIysAspAlaIysAspPheGlyIleLys 220  
Db 2627 AAGAAAGAACTTGAAGGTACACAAAATTTAGAGAAGCTAAAGATTTGGATAAAG 2686  
Qy 221 ArgThrIleAlaSerIysValSerLeuGlyAlaValTyPhePheMetAsnGlyThrTy 240  
Db 2687 AAGCTATTACAGCAATATTTCTATAGGTGCTCTTCTCTGCTGATCTATGATCATCTAT 2746  
Qy 241 GlyLeuAlaPheTrpTyThrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyThrIle 260  
Db 2747 GCTCTGGCCTTCTGTATGGACCACTTGGTCTCTCAGGGAA-----TATCTATT 2800  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyTyCysIleGlyAlaAla 280  
Db 2801 GGCAAGTACTCTACTGATTTCTTCTGTATTAATTGGGGCTTTTAGTGTGGACAGCA 2860  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 2861 TCTCCAAGCAATTGAAGCATTTGCAAAATGCAAGAGGACGCTTATGAAATCTTCAAGATA 2920  
Qy 301 IleAspIysLysProSerIleAspAsnPheSerThrAlaGlyTyIysProGluSerIle 320  
Db 2921 ATTGATTAATAGCCAGATTTTGACAGCTATTCGAAGTGGGCACAAACAGATAATT 2980  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyProSerArgProSerIleLys 340  
Db 2981 AAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 3040  
Qy 341 IleLeuIysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 3041 ATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGGACGCGTGGCCCTGTGTGGAAC 3100  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuGlnArgLeuTyArgProAspAsp 380  
Db 3101 AGTGGCTGTGGAGAGACACACAGTCCAGCTGATGACAGGCTCTATGACCCACAGAG 3160  
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyArgAsp 400  
Db 3161 GGGATGTCAGTGTGATGGACAGGATATTAGGACCATTAATGTAAGGTTTCTACGGAA 3220  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrIleSerAsnAsnIle 420  
Db 3221 ATCAATGGTGTGGTGCAGAACCTGTATGTGTGCCACGATAGCTGAAACATTT 3280  
Qy 421 LysTyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 3281 CGCTATGGCCGTGAATATGTCACCATGATGATGAGATGAGAAGCTGTCAAGAAAGCCAT 3340  
Qy 441 AlaTyArgPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460  
Db 3341 GCCTATGACTTTTATCATGAAACTGCCTCATAAATTTGACACCCCTGGTTGGAGAGAGGG 3400  
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 3401 GCCCAGTTAGTGTGGGCAAGACAGAGGATCGCCATTGACGTCGCCCTGTGTTCGAAC 3460  
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 3461 CCCAAGATCTCTCTGCTGGATGAGGCCACGTCAGCCCTTGGACAGAAAGCCAAAGCATG 3520  
Qy 501 ValGlnAlaAlaLeuGluLys 507  
Db 3521 GTTCAGGTGGCTCTGGATAAG 3541

## RESULT 7

US-08-793-610-6  
; Sequence 6, Application US/08793610  
; Patent No. 5858744  
; GENERAL INFORMATION:  
; APPLICANT: BAUM, Christopher  
; APPLICANT: STOCKING-HARBERS, Carol  
; APPLICANT: OSTERTAG, Wolfram  
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
; TITLE OF INVENTION: FOR GENE TRANSFER  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Nikaido, Marmelstein, Murray & Oram LLP  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,610  
; FILING DATE: 07-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 31 973.8  
; FILING DATE: 08-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 03 952.1  
; FILING DATE: 07-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/03175  
; FILING DATE: 10-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berman, Richard J.  
; REGISTRATION NUMBER: 39,105  
; REFERENCE/DOCKET NUMBER: P1614-7007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9318 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA  
US-08-793-610-6  
Alignment Scores:  
Pred. No.: 5,85e-178 Length: 9318  
Score: 1570.00 Matches: 305  
Percent Similarity: 76.85% Conservative: 100  
Best Local Similarity: 57.87% Mismatches: 100  
Query Match: 60.57% Indels: 22  
DB: 2 Gaps: 5  
US-09-873-409-8 (1-514) x US-08-793-610-6 (1-9318)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 1926 ATGGTGGTGGGAAGTCTTGGCTGCCATCATCCATGGGCTGGACTTCTCTCATGATGCTG 1985  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 1986 GTGTTGGAGAAATGACAGATATCTTTGCAAAATGCAGGAATTTAGAAAGATCTGATGCTCA 2045  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTySerPhePhe----- 43  
Db 2046 AACATCACTAATAGAAAGTGATATCAATGATACAGGGTTCTTTCATGAATCTGGAGGAAGAC 2105

Qy	44	-----ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr	61
Db	2106	ATGACCAGGTATGCCATTATTACAGTGGAAATGGTGTGGGTGCTGTTGTGCTGTAC	2165
Qy	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleAqlGlyGln	81
Db	2166	ATTCAGGTTTCATTTTGGTGCTCGAGCTGGAGACAAATACACAAATTTAGAAAAACG	2225
Qy	82	PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu	101
Db	2226	TTTTTTCATGCTATAATGCGACGAGGATAGCTGGTTGATGTCACGATGTTGGGAG	2285
Qy	102	LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle	120
Db	2286	CTTAACACCCGACTTACAGATGATGCTCTAAGATTAATGAAGTTATTGGTGACAAAT	2345
Qy	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys	140
Db	2346	GGAAATGTTCTTTCAGTCAATGGCACTATTTTCATCTGGGTTTATAGTAGGATTTACAG	2405
Qy	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla	160
Db	2406	GGTTGGAAGCTAAACCTTGTGATTTTGGCCATCAGTCTCTGTTCTTGGACTGTCAGCTG	2465
Qy	161	AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	2466	GTCGGGCAAGATACATCTCTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAAGCT	2525
Qy	181	GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	2526	GGACGAGTACTCGAAGGCTCTGGCAGCAATTAGACCTGTGTGATTCATTTGGAGGCAA	2585
Qy	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	2586	AAGAAAGAACTTCAAAAGGTACACAAAAAATTTAGAAAGAGCTTAAAGAATTTGGGATAAG	2645
Qy	221	ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr	240
Db	2646	AAAGCTATTACAGCCAATATTTCTATAGTGTGCTGTTCTCTGCTGATCATATGCAATCTAT	2705
Qy	241	GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle	260
Db	2706	GCTCTGGCCCTCTGGTATGGACCACTTGGTCTCTCAGGGAA-----TATCTATT	2759
Qy	261	GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla	280
Db	2760	GGACAAGTACTCACTGTATTCTTTTCTGTATTAAATTTGGGGCTTTTAGTGTGGACAGCA	2819
Qy	281	ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGluVal	300
Db	2820	TCTCCAAGCATTTGACATTTGCAATTCAGAGGAGCGAGCTTATGAATCTTCAAGATA	2879
Qy	301	IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle	320
Db	2880	ATTGATAATAAGCCAAGTATTGACAGCTATTTCGAAGAGTGGGACAAACCCAGATAAATATT	2939
Qy	321	GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys	340
Db	2940	AAGGGAAATTTGGAATTCGAAATGTTCACCTTTCAGTTACCCCATCTCGAAAAAAGATTAAG	2999
Qy	341	IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu	360
Db	3000	ATCTTGAAGGCCCTGAACCTGAAGGTGACAGATGGGCAGACGGTGGCCCTCGTTGGAAAC	3059
Qy	361	AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp	380
Db	3060	AGTGGCTGTGGGAAGACCAACAGTCCAGCTGATGTCAGAGGCTCTATGACCCACACAG	3119
Qy	381	GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp	400
Db	3120	GGGATGCTGAGTGTGATGGACAGGATATTAGGACCATTAATGTAAAGTTTCTACGGGAA	3179

Qy	401	HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle	420
Db	3180	ATCATTTGGTGTGTGAGTCAGGAACCTGTATTGTTGTGCCACCAGATAGCTGAAACATT	3239
Qy	421	LysTyrGlyIValArgAspValThrAspGluGluMetGluAArgAlaAlaAArgGluAlaAsn	440
Db	3240	CGCTATGGCGCGTGAANAATGTCCACATGGATGAGATTGAGAAGAGCTGTCAAGGAAGCCAA	3299
Qy	441	AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly	460
Db	3300	GCCTATGACCTTATCATGAACCTGCTCATATAAATTTGACACCTGGTTGGAGAGAGAGG	3359
Qy	461	AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAArgAlaLeuValArgAsn	480
Db	3360	GCCCAAGTTGAGTGTGGCGAGAGAGAGATCGCCATTGCACTGCGCCCTGCTGTCGCAAC	3419
Qy	481	ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla	500
Db	3420	CCCAAGATCTCTCTGCTGGATGAGGCCACGTGAGCTTGGACACAGAAAGCGAAGCAGTG	3479
Qy	501	ValGlnAlaAlaLeuGluLys 507	
Db	3480	GTTCAGGTGGCTCTGGATAAG 3500	
RESULT 8			
US-08-784-649A-1			
; Sequence 1, Application US/08784649A			
; Patent No. 5830697			
; GENERAL INFORMATION:			
; APPLICANT: Sikic, Branimir I			
; APPLICANT: Chen, Gang			
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO			
; TITLE OF INVENTION: CYCLOSPORIN MODULATION			
; NUMBER OF SEQUENCES: 5			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Fish & Richardson			
; STREET: 2200 Sand Hill Road			
; CITY: Menlo Park			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94025			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/784.649A			
; FILING DATE:			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Sherwood, Pamela J			
; REGISTRATION NUMBER: Reg.No. 5830697 36,677			
; REFERENCE/DOCKET NUMBER: 06037/007001			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415-322-5070			
; TELEFAX: 415-854-0875			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 4264 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: cDNA			
US-08-784-649A-1			
Alignment Scores:			
Pred. No.:		2,34e-178	Length: 4264
Score:		1568.50	Matches: 306
Percent Similarity:		77.04%	Conservative: 100
Best Local Similarity:		58.06%	Mismatches: 98
Query Match:		60.51%	Indels: 23
DB:			Gaps: 6

US-09-873-409-8 (1-514) x US-08-784-649A-1 (1-4264)

Qy 1 MetIleuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 292 ATGGTGGTGGAACTTTGGCTGGCCATCATCCATGGGCTGGACATTCCTCTCATGATGCTG 351  
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31  
Db 352 GTGTTTGGAGAAATGACAGATATCTTTGCAATGACAGAAATTTAGAGATCTGATGCTCA 411  
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43  
Db 412 AACATCACTAATAGAGATGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471  
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61  
Db 472 ATGACCAGGTATGCTTATTATACAGTGGAAATTTGGTCTGGGGTGGTGGTCTGTCTTAC 531  
Qy 62 IleGlnIleSerLeuTyrIleThrAlaAlaArgGlnThrLysArgIleAtqLysGln 81  
Db 532 ATTCAAGTTCATTTTGGCTGGCAGCTGGAGACAAATACACAAATTTAGAAACAG 591  
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101  
Db 592 TTTTTCATGCTATAATGCGACAGAGATAGCTGGTGGTGGTGGCAGATGTTGGGGAG 651  
Qy 102 LeuAsnThrArgMetThr-----AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120  
Db 652 CTTAAACCCGACCTTACAGATGATGTCTCAAGATTAATGAAGAAATTTGGTGACAAAT 711  
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140  
Db 712 GGA---ATGTTTCAGTCAATGGCAACATTTTCACTGGTATTAGTAGGATTTACAGT 768  
Qy 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160  
Db 769 GTTTGAAGCTAACCTTGTGATTTTGCCATCATGCTGCTGCTTCTTGGCAGTGTCTGCTGCT 828  
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180  
Db 829 GTCTGGCAAGATACCTATCTTCTTACTGATAAAGAACTCTTACGCTATGCAAAAGCT 888  
Qy 181 GlyAlaValAlaGluLeuValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200  
Db 889 GGACAGTAGCTGAAGAGCTTCTGGCAAAATTAAGCTGATGATGATGATGATGATGATGATGAT 948  
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220  
Db 949 AAGAAAGAACTTGAAGGTACACAAATAATTTAGAAAGCTAAAGAAATTTGGGATAAAG 1008  
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240  
Db 1009 AAAGCTATTACGCAATATTTCTATAGTGTGCTGCTTCTTCTGCTGATCTATGCTATCTTAT 1068  
Qy 241 GlyLeuAlaPheTyrTrpGlyThrSerLeuIleLeuAnGlyGluProGlyTyrThrIle 260  
Db 1069 GCTCTGGCTTCTGATGAGGACACCTTGTGCTCTCAGGGAA-----TATCTATT 1122  
Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280  
Db 1123 GGACAAAGTACTCAGTATTTCTTCTGATTAATTTGGGCTTTTGGTGGGACAGCA 1182  
Qy 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300  
Db 1183 TCTCCAAGCATTTGAAGCTTTGCAATGCAAGAGAGAGAGCTTATGAATAATCTTCAAGATA 1242  
Qy 301 IleAspLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320  
Db 1243 ATTGATAATAGCCCAAGTATTGACAGCTATTGCAAGAGTGGGACCAACCAAGATAATATT 1302  
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340  
Db 340

Db 1303 AAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAGTTAAG 1362  
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360  
Db 1363 ATCTTGAAGGCTTGAACCTGAAGGTGCAGGTGGCAGAGCGTGGCCCTGGTTGGAAC 1422  
Qy 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380  
Db 1423 AGTGGCTGTGGGAAGAGCACACAGCTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1482  
Qy 381 GlyPheIleMetValAspGluAsnAspIleAtqAlaLeuAsnValArgHisTyrArgAsp 400  
Db 1483 GGGATGGTCACTGTTGATGGACAGGATATTAGGACCATTAATGTAAGGTTTCTACGGGAA 1542  
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420  
Db 1543 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1602  
Qy 421 LysTyrGlyArgAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440  
Db 1603 CGCTATGGCGCTGAAATGTCACCATGATGAGATTGAGAAAGCTGCAAGGAAGCCAAAT 1662  
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460  
Db 1663 GCCTATGACTTTTATCATGAACTGCCTCATAAATTTGACACCTGTTGGAGAGAGGG 1722  
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480  
Db 1723 GCCAGTTGAGTGGTGGGAGAGAGAGAGAGATGCCATTCGACGTGGCTTGGTTCGCAAC 1782  
Qy 481 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500  
Db 1783 CCCAAGATCTCTGCTGGTGGATGAGGCCACCTGACCTCATAAATTTGACACCTGTTGGAGAGAG 1842  
Qy 501 ValGlnAlaLeuGluLys 507  
Db 1843 GTTCAGGTGGCTCTGGATAAG 1863

RESULT 9  
US-08-784-649A-5  
; Sequence 5, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; APPLICANT: Chen, Gang  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,649A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: Reg.No. 5830697 36,677  
; REFERENCE/DOCKET NUMBER: 06037/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5070  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 4264 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 US-08-784-649A-5

## Alignment Scores:

Pred. No.: 2,346-178 Length: 4264  
 Score: 1568.50 Matches: 306  
 Percent Similarity: 77.04% Conservative: 100  
 Best Local Similarity: 58.06% Mismatches: 98  
 Query Match: 60.51% Indels: 23  
 DB: 2 Gaps: 6

US-09-873-409-8 (1-514) x US-08-784-649A-5 (1-4264)

```

Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db 292 ATGGTGGTGGGAACCTTTGGCTGCATCATCCATGGGGCTGGACTCTCTCATGATGCTG 351
Qy 21 ValLeuGlyGluMetSerAsp-----AsnLeuIleSer 31
Db 352 GTGTTGGAGAAATGACAGATATCTTTGCAATGCGAGGAAATTTAGAGATCTGATGCTCA 411
Qy 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
Db 412 AACATCACTAATAGAGATGATATCAATGATACAGGGTCTTTCATGAATCTGGAGGAAGAC 471
Qy 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db 472 ATGACCAGATATGCCCTATTATTACAGTGAATTTGGTGGGGTGGCTGCTGTAC 531
Qy 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db 532 ATTCAGTTTCATTTTGGTGGCTGGCAGCTGGAGAACAAATACACAAAATTAGAAAACAG 591
Qy 82 PhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAspIleGlyGlu 101
Db 592 TTTTTCATGCTATATCGACAGGAGATAGGCTGGTTGATGTCACGATGTTGGGAG 651
Qy 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db 652 CTTAACACCCGACTACAGATGATGCTCCCAAGATTAAATGAGGAATTTGGTGCACAAAATT 711
Qy 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db 712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACCTGGGTTTATAGTAGGATTTACACGT 768
Qy 141 GlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db 769 GGTGGAAAGCTAACCTTGTGATTTGGCCATCAGTCTGTTCTTGGACCTGTCAGCTGCT 828
Qy 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db 829 GTCTGGGCAAGAGTACTATCTTCAATTTACTGATAAAGAACTCTTAGCTATGCAAAAGCT 888
Qy 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db 889 GGAGCAGTAGCTCAAGAGGCTTGGCAGCAATTTAGAACTGTGATTCATTTGGAGGACAA 948
Qy 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspPheGlyIleLys 220
Db 949 AAGAAGAACTTGAAGGTACACAAAATTTAGAGAGCTAAAGAAATTTGGATAAG 1008
Qy 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db 1009 AAAGCTATTACAGCAATATTTCTATAGGTGCTGTTTCTGCTGATCTATGATCTTAT 1068
Qy 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleAsnGlyGluProGlyTyrThrIle 260
Db 1069 GCTCTGCGCTTCTGGTATGGGACCACTTGGTCTCTCAGGGGAA-----TATTCTATT 1122
  
```

```

Qy 261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db 1123 GGCAAGTACTCACTGATTTCTTTCTGTTAAATTTGGGGCTTTTAGTGTGGCAGGCA 1182
Qy 281 ValProHisPheGlyThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db 1183 TCTCCAGCATTGAAGCATTTCGAAATGCAAGAGGAGCAGCTTATGAATCTTCAAGATA 1242
Qy 301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db 1243 ATTGATAATAAGCAAGTATTTCACAGCTATTTCGAAGAGTGGGCACAAACACAGATAATT 1302
Qy 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db 1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 1362
Qy 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db 1363 ATCTTGAAGGGCTGAACTGAAGTGCAGTGGCAGACGCTGGCCCTGGTTGGAAC 1422
Qy 361 AsnGlySerGlyLysSerThrValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db 1423 AGTGGCTGTGGGAAGACACACAGTCCAGCTGATGCAGAGGCTCTATGACCCACAGAG 1482
Qy 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db 1483 GGGATGTCAGTGTGATGCAGCAGGATATTAGGACCATAAATGTAAAGTTTCTACGGGAA 1542
Qy 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnIle 420
Db 1543 ATCAITGGTGTGTGAGTCAGAACCTGTATTGTTGGCCACACGATAGCTGAAACATT 1602
Qy 421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1603 CGCTATGCGCGTGAAATCTGCCATGGATGAGATTGAGAAAGCTCTCAAGGAAGCCAAT 1662
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1663 GCCTATGATCTTATCATGAAACTGCTCATAAATTTGACACCCCTGTTGGAGAGAGGG 1722
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1723 GCCCAGTTGAGTGTGGGAGAGCAGAGATCGCCATTGACGTCCTGTTGGTTGCAAC 1782
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysAla 500
Db 1783 CCCAAGATCTCTGCTGTGATGAGGCCACGTCAGCTTGGACACAGAAAGCAGCAGTG 1842
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1843 GTTCAGGTGGCTCTGGATAAG 1863
RESULT 10
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman, Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3:
; LENGTH: 4669
5206352-3
Alignment Scores:
  
```

```
Pred. No.: 3,14e-178 Length: 4669
Score: 1568.00 Matches: 305
Percent Similarity: 76.66% Conservative: 99
Best Local Similarity: 57.87% Mismatches: 101
Query Match: 60.49% Indels: 22
DB: Gaps: 5

US-09-873-409-8 (1-514) x 5206352-3 (1-4669)

QY 1 MetIleuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
DB 575 ATGGTGGTGGAACTTTGGTGGCCATCATCGGGCTGGACTTCTCTCATGATGCTG 634
QY 21 ValLeuGlyLeuMetSerAsp-----AsnLeuIleSer 31
DB 635 GTGTTGGAGAAATGACAGATATCTTTGCAATGCAGGAATTTAGAAATCTGATGTC 694
QY 32 GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe----- 43
DB 695 AACATCACTAATAGAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754
QY 44 -----ArgLeuThrLeuTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
DB 755 ATGACCAAGGTATGCTTATATACAGTGAATTTGGTGGTGGTGGTGGTGGTGGTGGT 814
QY 62 IleGlnIleSerLeuTrpIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
DB 815 ATTCAGGTTTCATTGGTGGCTGGAGCTGGAGACAAATACACAAAATTAGAAACAG 874
QY 82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
DB 875 TTTTTCATGCTATATCGACAGGAGATAGGCTGTTGATGTCACGATGTTGGGGAG 934
QY 102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
DB 935 CTTAAACACCGACTTACAGATGATGCTCTAAAGATTAATCAAGTTATTTGGTGAACAAAT 994
QY 121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
DB 995 GGAATGCTTTTCAGTCAATGGCAACATTTTTCAGTGGTTTATAGTAGATTTACAGCT 1054
QY 141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
DB 1055 GGTGGAAAGCTAACCTTGTGATTTGGCCATCAGTCTGCTTCTTGGACGTGCTGCTGT 1114
QY 161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
DB 1115 GTCTGGCAAGATATCTTCTTACTGATAAAGAACTCTTAGCGTATGCAGAAAGCT 1174
QY 181 GlyAlaValAlaGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
DB 1175 GGACGAGTCTGAGAGGCTTTGGAGCAATAGAACTGTGATGATTTGGAGGACAA 1234
QY 201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
DB 1235 AAGAAAGAACTTGAAGGTACACAAAATTTAGAAGAACTAAAGAAATTTGGGATAAG 1294
QY 221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
DB 1295 AAAGCTATTACAGCAATATTTCTATAGGTGCTCTTCTGCTGATCTATGATCATCTTAT 1354
QY 241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
DB 1355 GCTCTGGCTTCTGATGAGGACCACTTGGTCTCTCAGGGGAA-----TATCTATT 1408
QY 261 GlyThrValLeuAlaValPheSerValIleHisSerTyrCysIleGlyAlaAla 280
DB 1409 GGACAGTACTCACTGATTTCTTCTGATTAATTTGGGCTTTTAGTGTGGACAGCA 1468
QY 281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaPheHisIlePheGlnVal 300
DB 1469 TCTCCAAGCATTTGAAGCATTTGCAATGCAAGAGGAGCGCTTATGAATCTTCAAGATA 1528
```

```
QY 301 IleAspLysGlyProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
DB 1529 ATTGATAATAAGCAAGTATTGACAGCTATTGCAAGAGTGGGCACAAACCCAGATAATATT 1588
QY 321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
DB 1589 AAGGGAAATTTGGAAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGTTAAG 1648
QY 341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
DB 1649 ATCTTGAGGGCTGAACTGAAAGTGCAGTGGCAGACGGTGGCCCTGGTTGGAAC 1708
QY 361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
DB 1709 AGTGGCTGTGGGAAGAGCACACAGTCCAGTCCAGTCCAGAGGCTCTATGACCCACAGAG 1768
QY 381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
DB 1769 GGGATGCTGAGTGTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGAA 1828
QY 401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
DB 1829 ATCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1888
QY 421 LysTyrGlyArgAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440
DB 1889 CGCTATGGCGTGAATAATGTCACCATGATGATGATGATGATGATGATGATGATGATGAT 1948
QY 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlyLysGly 460
DB 1949 GCCTATGACTTTATCATGAAACTGCCTCAATAATTTGACACCTGTTGGAGAGAGAGGG 2008
QY 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
DB 2009 GCCAGTTGAGTGGTGGGAGAGAGAGAGAGATGCCATGCCATGCCATGCCATGCCATGCC 2068
QY 481 ProLysIleLeuLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
DB 2069 CCCAAGATCCTCTGCTGGATGAGGCCAGCTCAGCTCAGCTTGGACACAGAAAGCAGAGTG 2128
QY 501 ValGlnAlaLeuGluLys 507
DB 2129 GTTCAGGTGCTCTGGATAG 2149

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
```



/	PRIOR APPLICATION DATA:
/	APPLICATION NUMBER: 08/332,444
/	FILING DATE: 31-OCT-1994
/	APPLICATION NUMBER: 07/887,712
/	FILING DATE: 22-MAY-1992
/	INFORMATION FOR SEQ ID NO: 18:
/	SEQUENCE CHARACTERISTICS:
/	LENGTH: 4669 bases
/	TYPE: nucleic acid
/	STRANDEDNESS: singular
/	TOPOLOGY: linear
/	MOLECULE TYPE:
/	DESCRIPTION: Genomic DNA
/	US-08-583-276-18

  

Alignment Scores:			
Pred. No.:	6,15e-176	Length:	4669
Score:	1549.00	Matches:	301
Percent Similarity:	76.2%	Conservative:	101
Best Local Similarity:	57.1%	Mismatches:	103
Query Match:	59.7%	Indels:	22
DB:	2	Gaps:	5

  

US-09-873-409-8 (1-514) x US-08-583-276-18 (1-4669)			
QY	1	MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu	20
Db	575	ATGGTGTGGGAATTGGCTGCCATCATCCATGGGCTGGACTTCCTCTCATGATCGTG	634
QY	21	ValLeuGlyGluMetSerAsp-----AsnLeuIleSer	31
Db	635	GTCTTTGGAGAAATGCAGATATCTTTGCAATGCAGGAAATTTAGAAGATCTGATGTCA	694
QY	32	GlyCysLeuValGlnThr-----AsnThrTyrSerPhePhe-	43
Db	695	AACATCACCTAATAGAAAGTAGATCAATGATACAGGGTCTTCATGAATCTGGAGGAAGAC	754
QY	44	-----ArgLeuThrLeuTyrTrpValGlyIleGlyValAlaLaLeuIlePheGlyTyr	61
Db	755	ATGACCAGGTATGCCCTATTATTACAGTGGAAITGGTGTGGGTGCTGGTTGCTGCTTAC	814
QY	62	IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln	81
Db	815	ATTTCAGGTTTCATTTTTGGTGCCTGGCAGCTGGAACACAAATACACAATAATTAGAAAAACAG	874
QY	82	PhePheHiserValLeuAlaGlnAspileGlyTrpPheAspSerCysAspileGlyGlu	101
Db	875	TTTTTTCATGCTATAATGCGACAGGATAGCTGGTTTGATGTGCACGATGTTGGGCAG	934
QY	102	LeuAsnThrArgMetThr---AspileAspyIleSerAspGlyIleGlyAspLysile	120
Db	935	CTTAACACCGACTTACAGATGATGCTCTAAGATTAATGAATTAATTTGGTGACAAAATT	994
QY	121	AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVallys	140
Db	995	GGAAATGTTCTTTTCAGTCAAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT	1054
QY	141	GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaala	160
Db	1055	GGTTGGAAGCTAACCCCTGTGATTTTGGCCATCAGCTGCTGTTCTTGGAGCTGCTGAGCTGCT	1114
QY	161	AlaCysArserArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla	180
Db	1115	GTCTGGGCAAGATACATCTCTTCACTTACTGATAAGAACTCTTAGCGGTATGC AAAAGCT	1174
QY	181	GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln	200
Db	1175	GGAGCAGTAGCTGAAGAGGTCTTTGGCAGCAATTAGAACTGTGATTGCAATTTGGAGGACAA	1234
QY	201	GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys	220
Db	1235	AAGAAAGAACTTGAAAGGTAACAACAAAATAATTAGAAAGAGCTAAAAGAAATTTGGGATAAG	1294

STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/612,521  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hamilton, Amy E  
REGISTRATION NUMBER: 33,894  
REFERENCE/DOCKET NUMBER: X-9693  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3169  
TELEFAX: 317-276-1294  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4224 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4224  
US-08-612-521-1

## Alignment Scores:

Prod. No.: 7,87e-107 Length: 4224  
Score: 975.50 Matches: 209  
Percent Similarity: 59.81% Conservative: 114  
Best Local Similarity: 38.70% Mismatches: 180  
Query Match: 37.64% Indels: 37  
DB: 1 Gaps: 8

US-09-873-409-8 (1-514) x US-08-612-521-1 (1-4224)

QY 1 MetLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 466 ATGCTTGGGCTGTGCTGCGGTGACCGCGGCTTCCCAACCTTTAAATGACTTTA 525  
QY 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 526 ATATTCGGTTCGATTAAACCGCTCTTTACGAATTTATGCTGTCTATT---GCGAACCAATA 582  
QY 41 SerPhePheArgLeuThr----- 46  
Db 583 TCCCAAGGGGACTTACTCCGAGACGCTTCCGCGCTTTACAGCAGCAAGATGATCTC 642  
QY 47 -----LeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58  
Db 643 AAGACTCAATCTGSCCAATAGCCCTTTATCTGATGCGCATTTGSCATTTGGAATGTTTCTA 702  
QY 59 PheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLysArgIle 78  
Db 703 GCGACTTGGCTTTACATGTTTCATCTGGAATGTTACTGGCGGAGCTTCAACTCTAAAAGGATT 762  
QY 79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSerCysAsp 98  
Db 763 AGAGAGGCTTACTTGGCTGAGCTGTAGCAGGAGATTGCTTACTTTGATGATCTGGGC 822  
QY 99 IleGlyGluLeuAnthrArgMet---ThrAspIleAspLysIleSerAspGlyIleGly 117  
Db 823 GCGGAGAGTTCGCACTTCGCAATTCAGCCGATTTGTCACCTTTGTCACAGGAGGAACTCT 882  
QY 118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137  
Db 883 GAAAGGTCGCTCGTATTTCCAATATGCTGGTACTTTTGTCTGCGGTTTTGTTCTAGCT 942

QY 138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157  
Db 943 TTTGTCCGGTCACTCGTCTTGGGGCGGTCTTGTGTTTCCATCTCCAGTCATCATGCTC 1002  
QY 158 SerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177  
Db 1003 TGGCGCGGTATTATGATGACGGCCAAATATGGGACTCGGCTCTTGTATCATC 1062  
QY 178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197  
Db 1063 GCGAAGCGGCGAGCTTGGCAGAGAGGTCATAGGAGATATCAGGACTGTTTCAGGCTTT 1122  
QY 198 ArgAlaGlnGluLysGluLeu---GlnArgTyrThrGlnAsnLeuLysAspAlaLysAsp 216  
Db 1123 ---GGCAAGGAAAGATCTTGGGTGCAAAATTTGCCGATCACATTGAGCAGAGCAAGATT 1179  
QY 217 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMet 236  
Db 1180 GTCGGTAGAAAAGGCTCCATCTTTGAAGGCTTTGGTTTGAGCATCATGTTCTTCGTCATC 1239  
QY 237 AsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGluPro 256  
Db 1240 TAGCGCGTTATCCCTCGCTTCTTCTAGGTGGCATTCCTGTCAGTAAAGCCAGGCC 1299  
QY 257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCys 276  
Db 1300 -----GACTCCGGTATCGTCATCAACGTTTTCATGTCATCCTCATGTTCTTCTTCC 1353  
QY 277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 296  
Db 1354 ATGCGCATGCTTCTCCCGAGTTGGCTGTGTCCAAAGCGGAGCGGTGTGCCAAG 1413  
QY 297 IlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLys 316  
Db 1414 CTTTTCGCAACTATCAGCGGTGATCCCGCATTTGATTCGCCAGCGAGGAGGTTTCAAG 1473  
QY 317 ProGluSerIleGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 336  
Db 1474 CCGCATGGCTTCGCGGTGAGATTAGCTTTGAAAAATGTCAAGTTCCATTATCCGTCCGA 1533  
QY 337 ProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 356  
Db 1534 CCAGCATTCCTATCTTGAAGGCTTCTACTACTACGTTTGAAGCGGTAAAGACTTTGCT 1593  
QY 357 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyr 376  
Db 1594 CTGCTCGAGCTAGTGGAGCGGAAAGAGTACTGTCTGTTCTTCTTATGAGAGGTTTAC 1653  
QY 377 AspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 396  
Db 1654 GACCCGTTAGTGTGTGTGAAGCTTGAAGGTAGAGATATCAGGTCCGTAAACCTCAAT 1713  
QY 397 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrIle 416  
Db 1714 TGGCTTCGTCAACAGATTGCTCTGTATCTCAGGAACCCACTCTTTTCGGTACACCGCTC 1773  
QY 417 SerAsnAsnIleLysTyrGly-----ArgAspAspValThrAspGluGlu 431  
Db 1774 GCGGTAACGTCGAACACCGTCTCATCGGCTCCAGATATGAGAAATGCTCTCTCAGGAA 1833  
QY 432 -----MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGlu 447  
Db 1834 AAATTTGAGCTTGTGAAGAGGCTTGTGTCGATGCCAACGCCCAATAATTTATCATGAAA 1893  
QY 448 PheProAsnLysPheAsnThrLeuValGlyLysGlyAlaGlnMetSerGlyGlyGln 467  
Db 1894 CTTCCGCAAGGTTATGATACCATGTTGGTGGAGCGGTATGCTGTTATCTGTTGGTGCAG 1953  
QY 468 LysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAsp 487  
Db 1954 AAGCAGGAGTTCGATTGCTGCTGCTATCTGTTCCGACCTAGGATTTGTTGTTGGAC 2013

```

Qy 488 GluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGlnLys 507
|||
Db 2014 GAAGACTACCTCTGCTTGTGACACTCAGACGGAAGGTATTGTGCAAGATGCTTTGGACAAG 2073

RESULT 13
US-08-612-734B-1
; Sequence 1, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; TITLE OF INVENTION: Aspergillus Fumigatus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center, DCI501
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4047 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4047
; US-08-612-734B-1

Alignment Scores:
Pred. No.: 7,16e-105 Length: 4047
Score: 959.00 Matches: 203
Percent Similarity: 59.96% Conservative: 110
Best Local Similarity: 38.89% Mismatches: 191
Query Match: 37.00% Indels: 18
DB: 2 Gaps: 6

US-09-873-409-8 (1-514) x US-08-612-734B-1 (1-4047)
Qy 1 MetLeuGlyLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
|||||
Db 331 ATACTCGTGCGCAATCTGCTATTGCTGTGTCGGCTCTTCCCTTTGTTCAAAATT 390

Qy 21 ValLeuGlyGluMetSerAspAsnLeuLeuSerGlyCysLeuValGlnThrAsnThr 40
|||
Db 391 CTCPTCGGTTCACTAGCC---TCAGCGGTTCCAGGGCATATCTCTGGGTACTATGCCCTAT 447

Qy 41 SerPhePhe-----ArgLeuThr-----LeuTyrTyrValGlyIleGlyValAla 55
|||||
Db 448 CATGAGTTCTATCACAACTGACTAGAAATGCTTTACTTTGTGTATCTCGTATTCGCC 507

Qy 56 AlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThr 75

```

Db 1582 ACGATCTTTAGAAACATCGAACATGGCTTGATTTGGCCACCAAAATTCGAGCATAGTCGAAG 1641  
Qy 426 AspValThrAspGluMetGluAlaArgGluAlaAlaArgGluAlaAsnAlaTyrAspPheile 445  
Db 1642 GACAAGATTAGAGAGCTCGTTGAGAAATGCGGCCAGAAATGGCCATGCTCATGATTTATT 1701  
Qy 446 MetGluPheProAnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465  
Db 1702 ATGGCTCTGCCTGAAGGTTACGATACGAATGTGGGTGAGCGTGGTTCTTCTTCTTACGGA 1761  
Qy 466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuile 485  
Db 1762 GGTGAGAAGCAACGATTATGGCAATTCCTGTCGCCATTGTCACTGACCCCAAGATTCTGTTG 1821  
Qy 486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505  
Db 1822 CTTGATGAAGCTACATCAGCTTTGTGATACCAAGTCCGAGGCGTGCCTCCAGCGCTCTT 1881  
Qy 506 GluLys 507  
Db 1882 GATAAA 1887

## RESULT 14

US-08-612-734B-3  
; Sequence 3, Application US/08612734B  
; Patent No. 5914246  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: Tobin, Matthew B.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene of  
; TITLE OF INVENTION: Aspergillus Fumigatus  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center, DC1501  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.A.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,734B  
; FILING DATE: 08-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Craig, Anne I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: X-9681  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-354-9570  
; TELEFAX: 617-354-4043  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4800 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
US-08-612-734B-3

Alignment Scores:  
Pred. No.: 2.16e-102 Length: 4800  
Score: 939.50 Matches: 202  
Percent Similarity: 58.35% Conservative: 102  
Best Local Similarity: 38.77% Mismatches: 182  
Query Match: 36.25% Indels: 35  
DB: 2 Gaps: 6

US-09-873-409-8 (1-514) x US-08-612-734B-3 (1-4800)  
Qy 1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20  
Db 693 ATTCTCTTGGTTCATAGCTCAGCTCAGGCGATATCTCTGGGTACTATGCCCTAT 752  
Qy 21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40  
Db 753 -----CAT 755  
Qy 41 SerPhePhe---ArgLeuThr-----LeuTyrTyrValGlyIleGlyValAlaAala 56  
Db 756 GAGTTCTCACAACTGACTAAGAAATGTCTTTACTTTGTGTATCTCGGTATTCGGCAG 815  
Qy 57 LeuIlePheGlyTyrIleGlnIleSerLeuTyrIleIleThrAlaAlaArgGlnThrLys 76  
Db 816 TTTGTACAGTCTATGTACGACCGCTGGGTTTCATTTTACTGGCGAATCTCACACAG 875  
Qy 77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTyrPheAspSer 96  
Db 876 AAGATCCGTGAAATATTATCTTGAGGCTATCTCAGGCGAGAATATGGCTTACTTCGACAAG 935  
Qy 97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115  
Db 936 TTGGGCGCGCGTGAAGTTACACGCGTATATCCTGATACCAACCTGATCCAGGAGGCC 995  
Qy 116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135  
Db 996 ATCTCTGAGAAATTTGGTCTCCTACTTTGACCGCATTCGCCACATTTGTAACCGCATTTATT 1055  
Qy 136 ValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155  
Db 1056 GTCGCTACGTCAAGTATTGGAAGTTGGCTCTCATCTGTACCTCAACTATCGTCGCGCTG 1115  
Qy 156 MetAlaSerAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175  
Db 1116 GTCATGTTATGGGAGTGGTGGTTCAGGTTTATTGTGAAGTACAGCAAGAAATCTATTGAA 1175  
Qy 176 AlaTyrSerLysAlaGlyAlaValAlaGluValLeuSerSerIleArgThrValIle 195  
Db 1176 AGTTATGGTCTCGTGGAATGTCGCGGAAGAAGTATCATAGCTCCATTCGGAATGCTACC 1235  
Qy 196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215  
Db 1236 GCTTTCGGCACTCAGGATAAGCTCGCAAGCAATACGAACCCATCTGGCTGAGGCTGAA 1295  
Qy 216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235  
Db 1296 AAATGGGCGCTCAACAACAGAGTCTATCTGGTATGATGATGTTGGTGTATGTTCTGGTATC 1355  
Qy 236 MetAsnGlyThrTyrGlyLeuAlaPheTyrTyrGlyThrSerLeuIleLeuAsnGlyGlu 255  
Db 1356 ATGTTCTCGAACTATGTCGCTCGGTTCTGGATGGGATCTCGCTTGTGTGCGTAAGAA 1415  
Qy 256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275  
Db 1416 -----GTCAACCTGGGCCAAGTTCTGACAGTTTGTATGTCATCTCGATCGGTTCGTT 1469  
Qy 276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295  
Db 1470 AGTTGGGCAACGTCGCCCCCAATGTCAGGCGCTTTACGAATGGTGTGTCGGCGCGCG 1529  
Qy 296 HisIlePheGlnValIleAspLysLysProSerIleAsnAsnPheSerThrAlaGlyTyr 315  
Db 1530 AAGATTTACAGCGCATTTGACCGCAGATCGCCACTGGACCCCTATTCTGACGAAGGAAG 1589  
Qy 316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335  
Db 1590 GTACTCGACCATTTTGAAGGAATATCGAATTTGCAATGTCGAACACATCTACCCCTCA 1649  
Qy 336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355  
Db 1650 AGACCCGAAGTTACAGTCAAGTGAAGATGCTCTTTATCGATGCCGCGCGGAAGACTACC 1709

```
QY 356 AlaLeuValGlyLeuAenGlySerGlyLysSerThrValValGlnLeuGlnArgLeu 375
D 1710 GCATTGGTGGCCCATCTGGCTCTGGAAGAGTACTGTTGTCGCTTGGTGGAGCGCTTT 1769
QY 376 TyrAspProAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAenVal 395
D 1770 TACCTTCCAGTAGGAGCGCAGGTATTGCTGGACGGCCATGATATCCAAACCCCTCAACCTC 1829
QY 396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
D 1830 CGTTGGCTCGACAGCAGATCTCTCTGTGTGACGCCAGGACCTCTCTTTTCAGCACCCAG 1889
QY 416 IleSerAsnAsnIleLysTyrGly-----ArgAspAsp 426
D 1890 ATCTTTAGAACATCAACATGCTTGAATGGCACCACAAATTCGAGCATGATCGAAGGAC 1949
QY 427 ValThrAspGluMetGluArgAlaAlaAargGluAlaAsnAlaTyrAspPheIleMet 446
D 1950 AAGATTAGAGAGCTCGTTGAGAAATGGGCCAGAAATGGCCAAATGCTCATGATTTTATTATG 2009
QY 447 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly 466
D 2010 GCTCTCCCTGAAGTTACATACATGTTGGTTCAGCGTGGTTTCTTACTTTCAGGAGGT 2069
QY 467 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 486
D 2070 CAGAAGCAACGATTGCTGCTGTCGTCATGTCAGTGACCCCAAGATTCCTGTTGCTT 2129
QY 487 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaLeuGlu 506
D 2130 GATGAAGCTACATCAGCTTTGGATACCAAGTCGAGGGCGCTGCTCAAGCGCTCTTGAT 2189
QY 507 Lys 507
D 2190 AAA 2192
RESULT 15
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1
```

```
Alignment Scores:
Pred. No.: 2,75e-102 Length: 4002
Score: 937.50 Matches: 201
Percent Similarity: 59.96% Conservative: 103
Best Local Similarity: 39.64% Mismatches: 181
Query Match: 36.17% Indels: 23
DB: 2 Gaps: 5
```

US-09-873-409-8 (1-514) x US-08-996-545-1 (1-4002)

```
QY 15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
D 368 GTGCCATTGCTGCGCGTCGACTTTCCAGA-----GGATAATGTTAT 409
QY 35 ValGlnThrAsnThrTyrSerPhePhe-----ArgLeuThrLeuTyrTyrVal 50
D 410 ATCAATCTCGTACGACGACGAGTTGTCTATGATGTAATTGACCAAGACGCTACTGTTCTGTA 468
QY 51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuThrIleThr 70
D 469 TACCTCGGTATCGCGAGTTGTCTACTGTCTATGTTAGTACTGTTGCTTCTCATCTATACC 528
QY 71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
D 529 GGAGAACACGCCACGAGAGATCCGCGAGTATTACCTTTCAGTCTATCTCTGCGCCAGAAC 588
QY 91 IleGlyTyrPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
D 589 ATTGGCTATTTTGTAACTCGTGGCGGGAAGTACCACCCGTTATACAGCCGATACA 648
QY 110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
D 649 AACCTTATCCAGGATGCGATTTCGGAGAAGGTGCGTCTCACTTTGACTGCGCTGGCGACA 708
QY 130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTyrLysLeuThrLeuValThrLeu 149
D 709 TTCGTGACAGCATTCATTATCGCTACGTCAAATACTGGAAGTTGGCTCTCTAATTTCGAGC 768
QY 150 SerThrSerProLeuIleMetAlaSerAlaAlaCysSerArgMetValIleSerLeu 169
D 769 TCAACAAATTGTGGCCCTCTCTCACCATGGGCGGTGTTCTCAGTTTATCATCAAGTAC 828
QY 170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
D 829 AGCAAAAGTCTGTTGACAGCTACGGTGCAGCGGCGACTGTTGCGGAAGAGGTTCATCAGC 888
QY 190 SerIleArgThrValIleAlaPheArgAlaGlnGlnLysGluLeuGlnArgTyrThrGln 209
D 889 TCCATCAGAAATGCCACAGGTTTGGCACCAAGAGCTTGGCAAGAGTATGAGGTC 948
QY 210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
D 949 CACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAATGTCATGGGTTTCATGATT 1008
QY 230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTyrTyrThrSer 249
D 1009 GGGCCATGTTGGCCCTTATGACTCGAAGTACGGTCTTGGCTTCTGGATGGGTTCTCGT 1068
QY 250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
```

```
Db 1069 TTCTCGTAGATGGT-----GCAGTCGATGTGGGTGATATTCTCACAGTTCTCATGGCC 1122
Qy 270 ValIleHisSerTyrCysIleGlyAlaValProHisPheGluThrPheAlaIle 289
Db 1123 ATCTTGATCGGATCGTTCTCTTGGGAACGTTAGTCCAAATGCTCAAGCATTTACAAAC 1182
Qy 290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
Db 1183 GCTGTGGCGCGCGCAAGATATTGGAAACGATCGATCGCCAGTCCCATAGATCCA 1242
Qy 310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db 1243 TATTGCAACGAAGGGAAGAGCGCTCGACCATTTTGGGGCCACATTTAGTTACGCAATGTC 1302
Qy 330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db 1303 AAGCATATTTACCCATCTAGACCCGAGGTACCGTCAAGGATGTTTCTGTCAATG 1362
Qy 350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db 1363 CCGCTGGAAAAACAACCGCTTTTAGTCGGCCCTCTGGCTCTGGAAAAAGTACGGTGGTC 1422
Qy 370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
Db 1423 GGCTTGTTGAGCGATTCTACATGCCCTGTTGCGGGTACGGTTTGTGGATGGCCATGAC 1482
Qy 390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
Db 1483 ATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCTTTGGTTAGCCAGGACCT 1542
Qy 410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly----- 423
Db 1543 GTTCTTTTGGCAGCAGATTTATAAGATATTAGGCACGGTCTCATCGCACAAAGTAC 1602
Qy 424 -----ArgAspAspValThrAspGluMetGluArgAlaAlaArgGluAlaAsn 440
Db 1603 GAGAAATCAATCCGAGGATAAGGTCCGGGAATCATCGAGAACCGGCAAAATGGCGAAT 1662
Qy 441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db 1663 GCTCATGACTTTATTACTGCCCTTCCTGAAGGTTATGAGACCAATGTGGCGACGGTGGC 1722
Qy 461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db 1723 TTTCTCTCTTCAGGTGGCCAGAAACAGGCATTGCAATCGCCCGTGGCGTGTGTAGTGAC 1782
Qy 481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db 1783 CCAAAAATCTCTGCTCTGGATGAAGTACTTTCGGCCTTGGACACAAAATCCGAAGCGGTG 1842
Qy 501 ValGlnAlaAlaLeuGluLys 507
Db 1843 GTTCAAGCAGCTTTGGAGAGG 1863
```

Search completed: March 31, 2003, 15:38:51  
Job time : 138.273 secs